D0.52
NAS8-36955

REAL-TIME QUALITY ASSURANCE TESTING USING PHOTONIC TECHNIQUES

APPLICATION TO IODINE WATER SYSTEM

JRC Research Report # 90-28

The University of Alabama in Huntsville

REAL-TIME QUALITY ASSURANCE TESTING USING PHOTONIC TECHNIQUES

APPLICATION TO IODINE WATER SYSTEM

JRC Research Report # 90-28

NASA George C. Marshall Space Flight Center
Huntsville, AL 35812
Contract No. NAS8-36955 D. O. 52

July 1990

Report prepared by

W. F. Arendale, Richard Hatcher, Yadilett Garlington, Jack Harwell, and Tracey Everett

LABORATORY FOR INLINE PROCESS ANALYSES
KENNETH E. JOHNSON RESEARCH CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

REAL-TIME QUALITY ASSURANCE TESTING USING PHOTONIC TECHNIQUES

APPLICATION TO IODINE WATER SYSTEM

ABSTRACT

A feasibility study of the use of inspection systems incorporating photonic sensors and multivariate analyses to provide an instrumentation system that in real-time assures quality and that the system is in control has been conducted. A system is in control when the near future of the product quality is predictable. Off-line chemical analyses can be used for a chemical process when slow kinetics allows time to take a sample to the laboratory and the system provides a recovery mechanism that returns the system to statistical control without intervention of the operator. The objective for this study has been the implementation of do-it-right-the-first-time and just-in-time philosophies

The ECLSS water reclamation system that adds iodine for biocidal control is an ideal candidate for the study and implementation of do-it-right-the-first-time technologies.

REAL TIME QUALITY ASSURANCE TESTING USING PHOTONIC TECHNIQUES

APPLICATION TO THE IODINE WATER SYSTEM

TABLE OF CONTENTS

TABLE OF FIGURES

REAL TIME QUALITY ASSURANCE TESTING USING PHOTONIC TECHNIQUES APPLICATIONS TO THE IODINE WATER SYSTEM

I. INTRODUCTION

The Laboratory for Inline Process Analyses of the Kenneth E. Johnson Research Center at The University of Alabama in Huntsville under Contract NAS-36955 D.O. 52 is conducting a feasibility study on inspection systems incorporating photonic sensors and multivariate analyses to be used to assure quality. The objectives are the implementation of do-it-right-the-first-time and just-in-time philosophies. Quality assurance monitoring using fiber optic techniques with spectroscopic analyses is an emerging technology. Combinations of optical fibers and instrumentation systems are being assessed for applicability to candidate materials and processes of interest to NASA Marshall. Demonstrations are selected through discussions between Drs. Ray Clinton and Richard Congo, NASA Project Monitors, and the UAH Principal Investigator, Dr. W. F. Arendale. It is anticipated that several sensor- instrumentation-fiber combinations may exhibit potential for characterization of materials of interest to NASA and associated contractors. Priority in selecting demonstrations has been given to applications of immediate interest to significant sized groups of NASA Marshall personnel.

II. SELECTION OF IODINE ANALYSES

The Principal Investigator was introduced to NASA personnel working on the ECLSS water reclamation system by Dr. Congo. Two areas were identified where real-time and multivariate analyses should offer immediate advantages over current procedures:

1. Identification of QA/QC process control variables useful on the water reclamation system. The controls should be inline to assure that the desired quality is available at all times and to provide a warning if the system is jolted from a statistically stable system.
2. Measurement of the rate of diffusion of iodine through teflon.

IDENTIFICATION AND INTERPRETATION OF CONTROL VARIABLES

Iodine in aqueous solution can exist as HI, I^-, $I_2(aq)$, I_3^-, HIO, OI^-, HIO_3, IO_3^-, HIO_4, IO_4^- and/or H_5IO_6. These species represent oxidation numbers for iodine of -1 to +7. These species compete with the H_2O, dissolved O_2, and dissolved H_2 for electrons. Further competition for the electrons occurs when organic species are present. The redox sensitivity is believed to be responsible for many conflicts related to analytical procedures. The water reclamation system will contain organics when metabolic degradation products are present and/or chemicals are used for experiments, cleaning or other operations.

A. MANY E° MUST BE CONSIDERED

The E° values for the Iodine - water system and a few typical organic species are shown in Table 1. The equations in Table 1 can be combined with equilibrium constants for known reactions to give the following equations that have been used in the literature as models for specific observations:

$$I_2(aq) + I^- \quad <----> \quad I_3^- \qquad K = 710 \qquad (1)$$

$$I_2(aq) + H_2O \quad <----> \quad H_2OI^+ + I^- \qquad K = 3 \times 10^{-13} \qquad (2)$$

$$I_2(aq) + H_2O \quad <----> \quad H^+ + I^- + HIO \qquad K = 5.4 \times 10^{-13} \qquad (3)$$

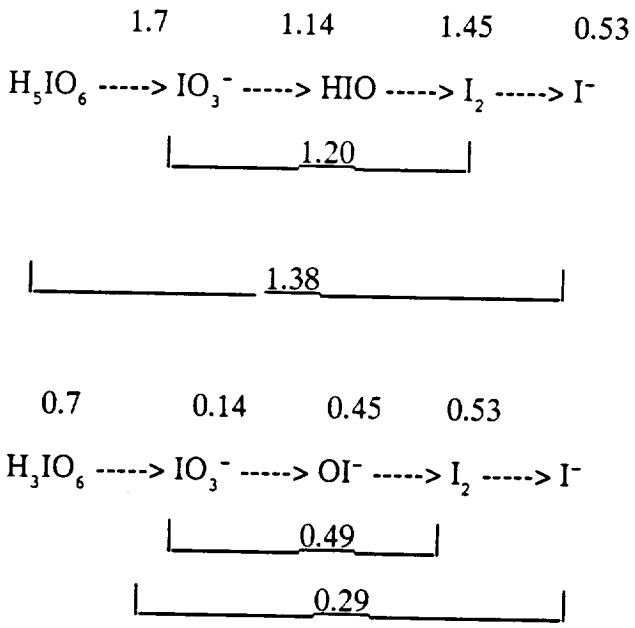
We have the possibility of 11 species related to iodine chemistry being found in the solution simultaneously. Therefore, for iodine species, there must be a minimum of eleven observations if we wish to know the thermodynamic state of the system.

Iodide in water is thermodynamically unstable with respect to dissolved oxygen.

$$4H^+ + O_2 + 6I^- \quad <----> \quad 2I^-_3 + 2H_2O \qquad (4)$$

This reaction is slow but is catalyzed by light, metal ions and organic species. Slow kinetics also keeps some species observable as redox equilibrium is not immediately reached. The concentration of each species also shifts as the result of changes in pH, pO_2, pCO_2, the presence of metal ions, and reaction with organic species (TOC).

Latimer Diagram



ELECTRODE	REACTION	*E°	E°'
	$CO_2(g) + 8H^+ + 8e$ <-----> $CH_4(g) + 2H_2O$	.17	0.24
$I_2/I^-/Pt$	$I_2(aq) + 2e$ <-----> $2\,I^-$	0.536	0.536
	$CH_3OH + 2H^+ + 2e$ <-----> $CH_4(g) + H_2O$	0.59	0.17
$O_2/H_2O_2/Pt$	$O_2 + 2\,H^+ + 2e$ <-----> H_2O_2	0.69	0.295
	$2H^+ + 2I^- + 1/2\,O_2$ <-----> $I_2(aq) + H_2O$	0.69	
	$NO_3^- + 10H^+ + 8e$ <-----> $NH_4^+ + 3H_2O$	0.88	0.36
$IO_3\text{-}/IO_3^-/Pt$	$12H^+ + 2IO_3^- + 10e$ <-----> $I_2(aq) + 6\,H_2O$	1.20	
$O_2/H_2O/Pt$	$O_2 + 4\,H^+ + 4e$ <-----> $2\,H_2O$	1.229	0.816
	$2NO_3^- + 12H^+ + 10e$ <-----> $N_2(g) + 6H_2O$	1.25	0.748
$HIO/I_2/PE$	$2H^+ + 2HIO + 4e$ <-----> $I_2(aq) + 2H_2O$	1.45	

*E° is given for all species at 1 molar concentration. Biochemist work with dilute aqueous solutions at pH near 7. E°' is E° adjusted from pH 0 to pH7.

Table 1.

When considering dilute solutions of I_2 in conductivity water where disproportionation, equation (3), is postulated, the quantity of H^+ is small and may be of the same magnitude as the H^+ from the reaction

$$H_2O <-----> H^+ + OH^- \quad (5)$$

The amount of H^+ from either reaction may not equal the of H^+ derived from an aqueous solution saturated with carbon dioxide.

$$CO_2(aq) + H_2O ----> H^+ + HCO_3^- \quad (6)$$

Examples have been found in the literature where dilute iodine solutions have been analyzed without protecting the sample from the carbon dioxide in the atmosphere. This is a serious mistake if the iodine concentration is only a few ppm, species are changing, and univariate analytical methods are being used. However open systems are standard practice when the concentration of iodine is 0.5F. This observation may account for the confusion concerning the chemistry of dilute aqueous iodine solutions and the variance in the analyses as performed by different organization and operators.

The reduction

$$I_2 + 2e <----> 2I^- \quad (7)$$

does not contain H^+ in the balanced equation. Therefore, E and the equilibrium constant K do not change with the concentration of H^+. The equation

$$I_2 + 2H_2O <----> 2HIO + 2H^+ + 2e \quad (8)$$

contains $2H^+$. Therefore, E° and K are functions of the square of the concentration of H^+, $[H^+]^2$. As acidity decreases, pH increases and more HIO is formed. As acidity increases, the equilibrium shifts to the left and I_2 or I_3^- is formed. As the concentration of HIO and I_3^- increase the following reaction becomes important.

$$3\ HIO + 3OH^- <----> IO_3^- + 2I^- + 3H_2O \quad (9)$$

The IO_3^- is a stable species. Species on the right are observed in open systems beginning at pH7 and the reaction goes to completion to the right at pH 10. At pH1 the reaction is quantitative to the left and as reaction (8) shows I_2 (aq) or I_3^- is the product. This model is the basis for one of the best methods for making standard solutions of iodine.

Determination of all species simultaneously and in real time is possible, highly desirable, and most likely the only way that the appropriate iodine species can be controlled. The concentration of each species is immediately available and there is no opportunity for the concentration of species to change while the samples await analyses. Should the sum of the iodine for each species in a closed system decrease then iodine has probably added to an organic species. The requirements for a successful controlled system are related to the requirements for solving simultaneous equations. A single unknown quantity can be evaluated using one equation. To solve for two unknowns a minimum of two independent equations is required. To solve for ten unknowns requires a minimum of ten independent equations.

B. PHYSICAL TOOLS AVAILABLE

The apparatus shown on Figure 1 is designed to evaluate the minimum number of sensors that will be required to determine that a stable process is operational. A system that is in statistical control has a definable identity and a definable capability. It is a random process. Its behavior in the near future is statistically predictable in the absence of unforeseen actions that knock the process out of statistical control. The apparatus shown on Figure 1 is designed to be used for experiments and simultaneously identify the statistical variation as measured by many different sensors exposed to a well defined system. The capability to jolt or upset the system and follow return to a stable system is also provided. After the system has been calibrated, the apparatus can be used to simulate full scale equipment when engineering experiments are required for design or redesign of the system. The following features are provided:

1. Charging port The tube and two way stopcock at the lower left is used to fill the system. To fill the system with water that does not contain dissolved oxygen, the system is flushed with an inert gas. Nitrogen can be

used if nitrogen fixation is not being investigated. Helium or argon can be used if both oxygen and nitrogen must be excluded. When the apparatus is free of active gases, then the sample bottle is pressurized and the apparatus filled. Both the two-way and three way stopcocks are then closed.

2. Septum The arms attached to the three way stopcock can be closed with a septum. Samples may be taken or reagents added using a syringe and needle placed through the septum. This port can be used to adjust pH by adding acid or base. Carbon dioxide can be added when needed to simulate the carbonate decomposition cycle or adjust acidity. Oxygen can be added as required to adjust oxidation potential. This port will be useful when studying the behavior of the solution after an upset in conditions. The port can also be used to take samples without exposing the sample to additional stimuli.

3. Test section The glass tubing in either leg of the system can be replaced with active devices. The tubing is the same diameter and length of a MCV. Filters could also be used.

4. Impeller The solution can be circulated using a variable speed motor.

5. Sensor ports Two ports are shown. These ports are used to incorporate sensors. Sensors currently used include pH, pI^-, pO_2, pCO_2, conductivity and/or pE.

6. Spectroscopic windows Five centimeter path lengths are provided for spectroscopic observation. UV grade quartz, NIR grade quartz, and sapphire windows are available. A window at right angles to the UV path will be added for observation of fluorescence. Additional path lengths will be added on the next design. Absorption probes with 1 and 2 cm path lengths are available and can be placed through the sensor ports.

7. Spectrophotometer A fiber optic cable is used to connect the light source to one window and a second fiber cable completes the path to the

spectrophotometer. Detectors include a photomultiplier, silicon, germanium, and lead sulfide. Grating available include 1200, 600, and 300 l/mm. This enables observation of high resolution spectra from 190 nm to 2200 nm. This includes the UV, Visible, and NIR regions of the electromagnetic spectrum.

8. Additional sensors An ATR sensor constructed from siloxane coated fiber is being used to observe the spectra for organophilic - hydrophobic species. This sensor is shown on Figure 2. Data obtained using this sensor is given with the experimental data in a later section.

At least two sensors should respond to any species. If an acid is generated the event should be recorded as a change in level of signal for both the pH and conductivity sensors. If the species is organic one or more spectroscopic sensors should respond. Should biomass form within the system or deposit on the windows, it should be readily detected unless it forms in the same thickness on each type of material used in the construction. The fiber optic ATR cell requires diffusion of the organic material into the cladding. The NIR ATR absorption spectrum and the NIR absorption spectrum would not be expected to show the same response for high molecular weight species such as the biomass. Organic species containing active hydrogen such as hydroxyl, carboxyl or amine groups will show strong NIR absorption. Since hydrophilic groups will not diffuse into the siloxane cladding, the opportunity for differentiation is provided. Phenols, cresols etc. will have strong UV absorption. Iodine reacts with benzene type compounds such as phenols to give organic addition products and the removal of iodine from the system. The compounds containing the benzene ring can be quantitated in the UV.

The objective is to observe in real-time a sufficient number of signals that there is enough knowledge of the system to predict the behavior of the system in the near future. When

this condition is achieved the effort should be directed to an engineering study of conditions which cause an upset of the system.

All of the sensors described above are provided for in the current design. We do not have multiplexers or enough spectrophotometers to make all observations simultaneously.

C. MODELS FOR THE SYSTEM

Data is obtained as the records of human observations and the digitized output of many sensors. This raw data becomes useful information when it becomes related to the system observed. A highly useful technique is to relate the data to a model. For many individuals a model that simulates the process being monitored is the most useful. The minimum requirement of the model that we seek is that it take the database of past experiences with the system and data currently being gathered and determine if the process is in the control region where observed variations are normally distributed. The model should also predict at least the near future condition of the system and most importantly that product is satisfactory for its intended use.

Experiences from the past indicate that their are many types of models that can meet the desired requirements. Models can be classified into at least three types: (1) quantitative or analytical, (2) qualitative, and (3) heuristic or empirical. The characteristics of each type, with examples, are defined in Appendix A attached to this report.

During this feasibility study we have worked with several models:

1. Quantitative models In an earlier section chemical reactions, equilibrium constants, E values, and the effects of the Law of Mass Action were discussed. Attempts to formulate this knowledge into quantitative models that simulate the chemistry of the iodine water system have given insight into the absolute values that we can expect and the possible statistical distributions that can be expected from the chosen sensors.

A software package that we have used in the past incorporates equilibrium constants, redox potentials, and includes the capability to allow floating or fixed values for pH, acidity, oxygen level, and CO_2 level. An example of the results obtained are shown on Figures 3 and 4. On Figure 3 is shown the predicted concentrations of I_3^- and HIO, with the pE held at 10.51 (E = 0.62), as a function of pH when the solution is 1E-3F (260 ppm) and 1E-4F in iodine. The point where the lines for $[I_3^-]$ and [HIO] cross would be the pH that would be expected if the indicated amount of I_2 is dissolved in conductivity grade water that has been freed of dissolved oxygen and carbon dioxide and the equation

$$2\,I_2 + H_2O \text{<----> } HIO + H^+ + I_3^- \qquad (10)$$

is the equation being modeled. The concentrations are good approximatation to the real system. In the real world one equation can not define a condition that is changing while the concentration of all other species and other conditions are held constant. On Figure 4 is shown the results for pE of 9.01. Note for the data on Figure 4 part of the iodine is not represented. The pE of 9.01 (E = 0.53) is in a region of pE where other iodine species are present in observable concentrations. What can we learn from these simulations! (1) This model indicates that pH will be a highly desirable sensitive attribute to follow with control charts. From a knowledge of pH electrodes we can predict that the combined variance that is observed will have broader upper and lower control limits, UCL and LCL, on the control chart than is found for some systems, for example, those used to monitor the buffered effluent from a chemical manufacturing plant. (2) This sensitivity to pH can be turned into a most important asset as will be demonstrated later.

Least squares regression analyses were performed for six iodine (aqueous) samples. The analyses were done using the absorptivity values at 460nm and 285nm. The data is shown on Figures 5 and 6. Note that the regression coefficients are 0.9967 and

0.9691. However, .. is extremely disturbing that the observed points are not random, but rather are above the line on each end and below the line in the center. A model other than a straight line obtained by least squares regression is indicated. Textbook state that Beer's law (linear regression) does not describe a system where there are reacting species.

(2) Heuristic or Empirical Models During the past eight years we have investigated a number of methods for multivariate calibration and analyses. Among the techniques that have been investigated are least squares regression, principal component analysis, principal component regression, factor analysis, singular value decomposition, and partial least squares (PLS) (four different algorithms).

The model discussed in the previous section was a univariant model using Beer's Law

$$AU_\lambda = \varepsilon_\lambda CL \qquad (11)$$

where AU_λ is measured in absorption units at a single frequency, ε_λ is a molecular constant that is molecule and frequency dependent and the units for AU_λ are absorption per mole per centimeter path length, C is concentration in moles/liter, and L is path length in centimeters. When working at a constant path length ε_λ L is combined into a single constant K_λ. The model for this description of the system can be described using Figure 7 where three dimensions are represented. For a model for the analyses given in the previous section label λ_1 absorption at 460 nm. For the Figure three values have been placed on the axis. Six were used. λ_2 axis can be labeled absorption at 285nm. Three values have been placed on this axis. We have reported individually the statistical regression for the values observed at 460nm and 285nm. If the absorptivity at 450nm and 285nm represent the same compound by vector addition we will find a distribution along the axis λ_1 λ_2. The statistics for the distribution along this axis can now be computed. This process can continue by plotting the values for an additional wavelength along λ_3. The combined observations for a single

compound will plot along the λ_1 λ_2 λ_3 axis and then the statistical distribution can be computed for the combined observables.

We collected 1801 points between 190nm and 550nm. When all of these absorption measurements represent the same molecule they plot along a single axis in 1801 dimensional space.

On Figure 8 are shown the spectra obtained for six samples made by dissolving I_2 in water. Each spectrum is offset by a fixed constant. It appears that the variance is observed at related points in the spectra. Therefore, data reduction using all 1801 points was performed. The results are shown in Table 2. The axis along which the statistical variation was computed was selected by principal component analyses.

Data taken from our investigation to select the appropriate smoothing technique is shown on Figure 9. A 7 point Savitsky-Golay smoothed spectrum has been subtracted from a spectrum as obtained. Note that there is severe random noise at both ends of the spectrum where signal to noise ratio is low. The noise appears to be random and Gaussian distributed. When a single frequency analysis is performed, for example Figure 5 and 6, we use the combined signal plus noise observation that may be correct, low or high. We accept the variance and plot statistical control charts for the observation. We understand that we can not take corrective action based on this data. If the UCL and LCL are within the product specification we do not tamper with the system. Can we improve the observation by correcting for noise? Sure, take data at several frequencies on each side of the observation point and make a statistical fit to the data. This is exactly what is done when we use all 1801 points that we have taken and apply a principal component selection procedure. Each of the 1801 points can be considered for use as a control variable. 1800 interferants could be detected.

Analysis of d:\sc\data\NASI1.spc The calibration matrix is IODA.cal Prediction error: 1.16869 Component Concentration prediction IODINE(aq) .00000596	Predicted 5.96×10^{-6}	Actual 5.91×10^{-6}
Analysis of d:\sc\data\NASI2.spc The calibration matrix is IODA.cal Prediction error: .540592 Component Concentration prediction IODINE(aq) .00001193	 1.19×10^{-5}	 1.18×10^{-5}
Analysis of d:\sc\data\NASI3.spc The calibration matrix is IODA.cal Prediction error: .264006 Component Concentration prediction IODINE(aq) .00002296	 $2.29\ 10^{-5}$	 2.36×10^{-5}
Analysis of d:\sc\data\NASI4.spc The calibration matrix is IODA.cal Prediction error: .177944 Component Concentration prediction IODINE(aq) .00003573	 3.57×10^{-5}	 3.55×10^{-5}
Analysis of d:\sc\data\NASI5.spc The calibration matrix is IODA.cal Prediction error: 1.06696 Component Concentration prediction IODINE(aq) .00005945	 5.94×10^{-5}	 5.91×10^{-5}
Analysis of d:\sc\data\NASI6.spc The calibration matrix is IODA.cal Prediction error: 1.51757 Component Concentration prediction IODINE(aq) .00011785	 1.18×10^{-4}	 1.18×10^{-4}

Table 2

On Figure 10 the spectra for solutions of KI, I_2, and KIO_3 have been plotted. Now we find that we have added the variance resulting from the absorption of bonds and free electrons that form the molecules.

A factor analysis of this data in the 1801 dimensional space should show a minimum of 5 factors according to our knowledge of the system. We anticipate a factor representing at least HIO and I_3^- for which standard solutions can not be made in addition to the standards that were made. Since the factors representing these two compounds can be plotted, we have a method of determining the spectra of these molecules that we do not find on the storeroom shelf.

Multivariate calibration techniques are being investigated. The PLSplus routine, version 1, that is a part of the SpectraCalc software collection was used for this feasibility study. Iodine solutions (six levels), KI (five levels), KIO_3^-(five levels), and KI_3 (KI plus I_2 five levels) and six samples made as mixtures were used as calibration sets. The spectra were recorded as 1801 points between 190nm and 550nm . Results are shown on Tables 3 and 4. These results indicate that 1801 point spectra can be easily analyzed with good results. 1800 points is probably a significant overkill. We have not had time to optimize the selection of factors, their interpretation, or the number of observations required.

The calibration model

$$Y = F(X) \qquad (12)$$

that predicts y_i from the x-variables is determined from the information in data from a set of Calibration Objects (training set). It is the experimenters responsibility to insure that the Calibration Set contains enough information to solve the selectivity problems involved.

For very simple selectivity problems only a few calibration objects are needed. But in general the less the user knows about the object type being analyzed and the instrumentation being used, the more calibration data needed in order to insure a reliable prediction ability. The calibration objects must be representative for the population of objects from which Y are to be predicted later, both with regard to average quality and to variability. A sufficient VARIABILITY must also be included in the calibration objects: The level of EVERY major independent component or phenomenon, that may cause interference effects in future X-data, must vary in the Calibration Object Set. The actual level of these effects do not have to be known!

For most efficient calibration design, the user should thus select a set of Calibration Objects that:

(1) are typical

(2) span each of the individual phenomena that can be controlled.

(3) are randomly selected to ensure a chance of spanning the non-controllable phenomena.

Empirical models are robust. These models are based on the same types of statistics as used for univariate control charts. The techniques are easily understood in two or possibly three dimensions. The understanding that a computer may be programmed to operate in many dimensions is the knowledge that is new to many engineers.

(3) Qualitative Models We have discussed two extreme types of models as used for modeling process data. The quantitative model based on physical laws and mathematical equations, and the empirical model where we recognize that all measurements are subject to variation and use a Calibration Set to identify the sources of variation. When data is obtained from sensors that respond to molecular variation, we can identify chemical species. When modeling chemical processes, qualitative models can be conceived as a hybrid.

Analysis of d:\sc\data\NASI1.spc The calibration matrix is IODHA.cal Prediction error: 4.98557	Prediction	Actual
Component Concentration prediction IODINE(aq) .00000595 IODIDE .00000027 IO_3^- -.00000002	$.595 \times 10^{-5}$	$.595 \times 10^{5}$ (1.5 ppm)
Analysis of d:\sc\data\NASI2.spc The calibration matrix is IODHA.cal Prediction error: 2.48077		
Component Concentration prediction IODINE(aq) .00001113 IODIDE .00000058 IO_3^- .00000001	1.11×10^{-5}	1.18×10^{-5} (3 ppm)
Analysis of d:\sc\data\NASI3.spc The calibration matrix is IODHA.cal Prediction error: 1.32528		
Component Concentration prediction IODINE(aq) .00002449 IODIDE -.00000038 IO_3^- -.00000004	$.245 \times 10^{-4}$	$.236 \times 10^{-4}$ (6 ppm)
Analysis of d:\sc\data\NASI4.spc The calibration matrix is IODHA.cal Prediction error: 1.00855		
Component Concentration prediction IODINE(aq) .00003463 IODIDE .00000069 IO_3^- .00000004	$.346 \times 10^{-4}$	$.355 \times 10^{-4}$ (9 ppm)
Analysis of d:\sc\data\NASI5.spc The calibration matrix is IODHA.cal Prediction error: .444153		
Component Concentration prediction IODINE(aq) .0000592 IODIDE -.00000003 IO_3^- 7.253E-09	$.592 \times 10^{-4}$	$.591 \times 10^{-4}$ (15 ppm)
Analysis of d:\sc\data\NASI6.spc The calibration matrix is IODHA.cal Prediction error: 5.0481		
Component Concentration prediction IODINE(aq) .00010237 IODIDE -.00000479 IO_3^- .00000058	1.02×10^{-4}	1.18×10^{-4} (30 ppm)

Table 3
MULTIVARIATE CALIBRATION
STANDARDS I_2(aq), KI AND KIO_3

Analysis of d:\sc\data\NASI7.spc The calibration matrix is IODHA.cal Prediction error: 6.37383		Prediction	Actual
Component	Concentration prediction		
IODINE(aq)	.00000016		
IODIDE	.0000958		
IO_3^-	-8.073E-09	$.958 \times 10^{-5}$	1.0×10^{-5}
Analysis of d:\sc\data\NASI8.spc The calibration matrix is IODHA.cal			
Component	Concentration prediction		
IODINE(aq)	.00000021		
IODIDE	.00009948		
IO_3^-	.00000002	$.955 \times 10^{-4}$	1.0×10^{-4}
Analysis of d:\sc\data\NASI12.spc The calibration matrix is IODHA.cal Prediction error: 1.60365			
Component	Concentration prediction		
IODINE(aq)	.00000009		
IODIDE	.00000019		
IO_3^-	.00000998	$.998 \times 10^{-5}$	1.0×10^{-5}
Analysis of d:\sc\data\NASI13.spc The calibration matrix is IODHA.cal Prediction error: -.00000702			
Component	Concentration prediction		
IODINE(aq)	-.00000702		
IODIDE	.00002057		
IO_3^-	.00000101	1×10^{-6}	1.0×10^{-4}

Table 3 (continued)
MULTIVARIATE CALIBRATION
STANDARDS I_2(aq), KI AND KIO_3

Analysis of C:\sc\data\PP1.spc The calibration matrix is IODKA.cal Prediction error: 3.32626		Actual
Component	Concentration prediction	
IODINE	.98346	1.0 ppm
IODIDE	.00001107	1×10^{-5} F
IO_3^-	.00001035	1×10^{-5} F

Analysis of C:\sc\data\PP2.spc The calibration matrix is IODKA.cal Prediction error: 6.32323		
Component	Concentration prediction	
IODINE	1.98501	2 ppm
IODIDE	.00001015	1.0×10^{-5} F
IO_3^-	.00002008	2.0×10^{-5} F

Analysis of C:\sc\data\PP3.spc The calibration matrix is IODKA.cal Prediction error: 8.72592		
Component	Concentration prediction	
IODINE	2.95085	3 ppm
IODIDE	.0000216	1.0×10^{-5} F
IO_3^-	.00001059	1×10^{-5} F

Analysis of C:\sc\data\PP4.spc The calibration matrix is IODKA.cal Prediction error: 7.77016		
Component	Concentration prediction	
IODINE	2.02039	2 ppm
IODIDE	.00001714	2×10^{-5} F
IO_3^-	.00000903	1×10^{-5} F

Analysis of C:\sc\data\PP5.spc The calibration matrix is IODKA.cal Prediction error: 5.30173		
Component	Concentration prediction	
IODINE	3.02628	3 ppm
IODIDE	.00001011	1.0×10^{-5} F
IO_3^-	.00001	1×10^{-5} F

Analysis of C:\sc\data\PP6.spc The calibration matrix is IODKA.cal Prediction error: 22.2323		
Component	Concentration prediction	
IODINE	2.01594	2 ppm
IODIDE	.00002939	3×10^{-5} F
IO_3^-	.00001973	2×10^{-5} F

Table 4
MULTIVARIATE CALIBRATION
TRAINING SET MIXTURES OF I_2, KI, AND KIO_3

Analysis of C:\sc\data\PP7.spc
The calibration matrix is IODKA.cal
Prediction error: 26.4408

Component	Concentration prediction	
IODINE	1.01804	1 ppm
IODIDE	.0000305	3×10^{-5} F
IO_3^-	.00001018	1×10^{-5} F

Table 4 (cont.)
MULTIVARIATE CALIBRATION
TRAINING SET MIXTURES OF I_2, KI, AND KIO_3

From our quantitative model we know the direction that a second or third attribute should take as a result of an extreme variation in the first variable. If an increase in the first variable should result in a decrease in the values for the second and third attribute and this is observed, although the magnitude may be less than expected, then we do not tamper with the system. We know that actual quantities are subject to statistical variation. However if the value for all three increase through an expert system or using AI we would begin closer observation to prepare to take corrective action. For an example, the data indicates a small increase in pH will cause major change in HIO/I_3^- ratios [HIO] will increase. The opposite change occurs with a decrease in pH.

We have not attempted a demonstration for this type of model during this feasibility demonstration. This type of method would probably emerge as the final model.

III. DESCRIPTION OF EXPERIMENTS PERFORMED TO DETERMINE FEASIBILITY OF PHOTONIC-MULTIVARIATE ANALYSES SYSTEM

A. DIFFUSION OF IODINE THROUGH TEFLON

1. Design of Test Apparatus

A coating of Teflon lines the tanks for the water reclamation system. A sheet of Teflon is used as a bladder within the tanks. The questions presented was:

How fast will iodine diffuse through the layer of Teflon?

Ms. Mary Treweek provided samples of two types of Teflon in the form of approximately eighth inch thick by three inches in diameter sheets. The apparatus shown on Figure 11 is an adaptation of the apparatus described on Figure 1. A sample is mounted between two o-rings at the place indicated. A reduced pressure is created in a chamber below the Teflon specimen see Figure 11. The initial suggestion for an iodine test solution was to replace the indicated glass tubing section on Figure 1 with a MCV (microbial check valve) and accept the solution coming from the MCV, although iodine species are undefined, as being a typical solution from which iodine species might diffuse. Delays at Marshall in obtaining MCVs and delays in receiving parts from the glass blower resulted in alternate experiments to examine the diffusion process. Several tests produced results believed to be significant.

2. Experimental Results

a. First Experiment The apparatus shown on Figure 12 was used for the initial experiments related to iodine in a controlled system. When required the solution can be circulated using the variable speed peristaltic pump. Samples of flexible tubing required by the peristaltic pump and meeting FDA specifications and made from silicone, Viton fluoroelastomer, Tygon, and Norprene were obtained. The Viton tubing was used in the

initial assembly of the apparatus. The Guided Wave Model 200 spectrophotometer with fiber optic cables was used to observe the UV spectrum from 190 - 550 nanometers across the five centimeter optical cell. A 50 ppm iodine solution was placed in the apparatus and circulation initiated. The UV spectrum that was obtained is shown on Figure 13.

Examination of Figure 13 indicates that the usual band for iodine solutions at 450nm was not observed. The iodine was consumed before the entire spectrum could be recorded (5 minutes). The dominant feature in the spectrum is the unexpected appearance of an absorption band at 282nm. As additional spectra were obtained, the absorption band at 282 nm increased. An absorption band at 282 nm is characteristic of compounds containing a bezene ring. Analytical work including Raman spectrophotometry was done in an attempt to identify the compound. Since only extremely small amounts of material was collected, specific identification was not obtained.

All four types of the purchased tubing gave similar results. All tubing available contains this foreign material. The manufacturer was contacted, but he was unaware of the material, initially denying its existence. He latter called to state a sample of the material had been obtained. At last contact he had not established positive identification. This material is being used either as a lubricant during extrusion or is included in the tubing as a component of the plasticizer. The best guess is the latter postulate. The manufacture agreed with our finding that the material was in extremely small amounts, not enough to exceed FDA specifications. Reformulation would be required to eliminate the material. The experiment demonstrates the excellent response of the 5 cm UV cell to detect aromatic species in aqueous solutions in real-time.

A special tubing plasticized with mineral oil is used by a dog food manufacturer to place iodine in a dog food preparation. Use of this tubing would eliminate the interference at

282 nm. However, an organic would be present and we do not have a way to simultaneously monitor its presence in a solution containing iodine species until we obtain additional spectrophotometric equipment so that we can observe the UV and NIR regions of the spectrum at the same time.

The experiments confirm the response of the system and shows the value of inline analyses to obtain immediate results. The benzene derivative was unknown to the manufacturer and would probably have remained unobserved if the observations had not been made in real-time.

b. Results Observed Using Alternate Experimental Configurations The apparatus described on Figure 11 was assembled from components of the apparatus described in Figure 1. A sample of the translucent Teflon was used for the first experiment. A solution containing approximately 30 ppm iodine was placed above the sample. The space below was evacuated. The UV spectrum from 190 - 550 nm was observed daily for 14 days. No spectra were obtained that indicated iodine was diffusing through the membrane. A weak spectrum of a substance believed to be an organic containing a benzene ring was obtained. The experiment was terminated. When the Teflon test piece was examined it appeared to be covered with a tan stain where the test piece had been in contact with the iodine. The piece was white where the sample was covered by the o-ring used to assemble the apparatus. The tan stain was examined using surface spectroscopy techniques. The depth of the stained area is very thin and no iodine was identified. Possibly, the iodine has reacted with an impurity on the surface to form a tan dye that does not contain iodine.

A test piece selected from the opaque white Teflon samples was next assembled into the apparatus shown on Figure 11. When the apparatus was assembled for the second experiment, the pieces of glass tubing in the lower observation chamber did not touch leaving

silicone tubing exposed at both the upper and lower joints. The lower chamber was reduced in pressure. The silicone tubing sealing the joints was observed to be a very sensitive indicator for iodine. The tubing at the upper joint was stained after 24 hours and the color became darker each day of the 14 day test. The tubing at the lower joint did not show staining. The upper piece of tubing was a very efficient scavenger for iodine. When the apparatus was disassembled after 14 days no staining was observed on the test piece. The best postulate is that the sample is made from sintered teflon and the iodine found a path between the teflon particles forming the test piece.

c. Comments on Above Results All literature available to us indicates that the diffusion of iodine through Teflon is slow, very much slower than with other plastics where reaction probably occurs. Conversations with several investigators indicated that the properties of Teflon are highly dependent on the method of manufacture. For example, fused Teflon is much less porous than sintered Teflon. The surface can be very different from the interior bulk properties. Since the rate of diffusion through Teflon is slow, it would be desirable to obtain thinner specimens for use with our sensitive detection methods. Consideration was given to obtaining thin samples using a microtone; however, this would remove at least one surface. Obtaining thinner samples prepared by the manufacturer would give emphasis to the surface material. When the second sample showed the possible existence of a path through the material, no further work has been done on sample preparation as it is felt that samples of the actual material must be used. Since the results depend on the integrity of the sample, extreme care will have to be given to sample selection.

A sensitive indicator for measurement of iodine concentration is the coloration in the silicone tubing. The color change is rapid and a wide range of concentration can be covered. The color change could be quantitated by simple absorption or by specular

reflection absorption measurements. Ten days is probably as long as a sample of iodine will maintain a fixed concentration in the apparatus that was used of iodine. If longer observation periods are required the solutions will need to be replaced. If a silicone indicator is placed behind the Teflon and in contact with a transparent window, analyses could be performed without disturbing the indicator. This test would be simple enough that a large number of samples could be observed giving sufficient data for statistical significance.

An alternate experiment would be the use of multiple test apparatus as shown on Figure 1. To make this experiment cost effective, other engineering test objectives should be accomplished on the solution side of the specimen. This experimental design would have the disadvantage that the solutions will have different concentrations of iodine species. However, the apparatus would be realistic with respect to current plant designs.

B. FOURTEEN DAY STUDY OF IODINE AND WATER IN A CLOSED SYSTEM

The equipment was assembled as shown on Figure 1. For this experiment the spectrophotometer was operated with a deuterium lamp, 1200 l/mm grating and a photomultiplier detector. Electromagnetic radiation below 225 nanometers causes the formation of yellow color centers in UV grade quartz fibers. Our fiber cables show attenuation compared to new fibers. To shorten the light path the UV deuterium lamp was placed adjacent to the window. A radiation probe that uses a quartz lens to focus the received radiation onto a six bundle cable was used to connect the port to the spectrophotometer.

A 30ppm iodine solution was prepared using water from a Millipore filter. The iodine solution was forced into the system using nitrogen gas. Circulation was initiated and a spectrum from 190 to 550nm was obtained. Normally the computer would be programmed to take readings at fixed time intervals. Since this study was planned for a long

time period and we did not wish to leave the system exposed to the short wavelength UV radiation continuously the UV lamp was turned off between spectra collection. During this first extended run some difficulties were encountered. A window leaked and a bubble formed at the upper stopcock. It was not determined if the gas was all iodine. Since iodine has an intense color and this was not observed, it was probably a mixture of iodine and air. Although an effort was made to have tight glass-glass butt joints, near the end of the run slight coloration in the tubing was observed. The spectra are shown on Figure 14. Semiquantitative analysis were obtained by subtracting from each spectra the first spectrum multiplied by an appropriate constant and subtracting. The results are shown and plotted on Figure 15.

The derived spectra obtained by subtraction of constant times known spectrum are shown on Figure 16. The residuals between 190 and 350nm represent species other than I_2(aq)but these spectra are not easily recognized. The analysis was repeated using five fresh iodine samples. The residual spectra, Figure 17, are similar to one another but not readily recognizable.

The experiment will be repeated when we refined our calibration matrix.

C. EXPERIMENTS WITH FIBER ATR SENSOR

Under Air Force Contract FO8635-89-C-0288, we are working on ADVANCED FIBER OPTIC SENSORS. Some of the results are described in this section.

Recent experimental work has centered on the use of a coiled fiber optic three centimeters in diameter. A schematic is shown on Figure 2. The coil is wound from FIBER-GUIDE INDUSTRIES fiber APC400N. This fiber is a low hydroxy pure quartz fiber with a 400 micron core, 100 micron cladding, and 100 micron nylon jacket. After the coil

is wound, the jacket in the area used for sensing is removed by immersing the probe in boiling propylene glycol. Our design differs somewhat from the design reported by Degrandpre and Burgess, APPLIED SPECTROSCOPY, 1990, 44, 273. The fiber is threaded through small holes that have been drilled at the edges of the supports. The most immediate improvement is in useful life. The paper reports a useful life of two weeks. Our first sensor lasted six weeks. When the fiber is threaded through the holes, we achieve reproducibility in microstrain and hence reduction in signal variance. All measurements to date have been made without terminating the fibers with SMA 905 connectors. Better signal strength and reproducibility is expected with termination. We have sufficient signal strength that we are proceeding to investigate other operating parameters of the sensor without the expensive termination step.

A Guided Wave Model 200 spectrophotometer with a 600 l/mm grating and silicon and germanium detectors is being used to record the spectra in 1 nm increments between 800nm and 2000nm. A tungsten hydrogen lamp is used as the source of illumination.

The sensor is connected to the spectrophotometer by 0.75 meter fiber optic cable from which the nylon has not been removed. The ends of the fiber are cleaved with a sharp knife leaving a smooth optical surface perpendicular to the fiber. The fiber is mounted in position by a fiber clamp and adjusted with respect to the spectrophotometer optics by moving the fiber in and out until a maximum signal is obtained.

The signal is recorded as a digital signal. The raw data (without processing to remove noise) is plotted on Figures 18, 19, and 20. Figure 18 is the spectrum for a 1:1 mixture of ethanol and chloroform against a background of the signal obtained with the sensor in air. Since the Model 200 is a single beam spectrophotometer, a background spectra is required as absorption units, AU, are the logarithm of T/T°. The spectra are plotted in AU

vs NM. The spectra being obtained are ATR (attenuated total reflectance). Note that the AU axis has some negative numbers. The electromagnetic wave travels down the core by total reflection at the boundary of the core and cladding. An evanescent wave penetrates (microns) into the cladding. Total reflection does not occur when the cladding has an absorption band. Therefore, less light passed to the detector when the background was taken in those regions where the chemical bands in the polysiloxane absorb. When the fiber sensor is immersed in the sample, the organophilic-hydrophobic cladding absorbs organics. The absorbed material causes the cladding to swell. Therefore, there is less of the siloxane in the distance penetrated by the evanescent wave. These bands will be extremely useful during the planned multivariant calibration for the sensor. They will be used as internal standards and also perform the same function as measurement of cell thickness in standard absorption spectroscopy.

Figure 19 is the spectrum for hexane. The sensor is washed with acetone and allowed to dry before taking the spectrum of a different sample. Figure 20 shows the progressive uptake of toluene vapor (spectrum at thirty minute intervals) when the sensor was suspended above an aqueous solution saturated with toluene. The toluene source was removed. Figure 21 shows the spectrum after 15 minutes had elapsed. The transmission through the fiber has returned to normal. The sensor is easily overloaded with toluene.

The spectra of extremely dilute aqueous solutions of toluene can be obtained as shown on Figure 22 (concentration unknown but ppm).

Although the toluene or other organic is easily removed with acetone, the partition coefficient for toluene is so high that removal by continual leaching a saturated sensor with deionized water is difficult.

IV. COMMENTS AND CONCLUSIONS

A. CONTROL SENSORS FOR CURRENT ENGINEERING TEST

The schematic, available to us, of the current system shows sensors for conductivity, pH, redox potential and a single spectrophometric peak. The control variables that we have identified were provided for in the design of the system. Probably the only change would be in the design of the spectrophotometer. We have a preliminary a model for the redesign.

We propose that multivariate analyses should be made on the data that has been obtained during the current tests. (A data set that should be immediately examined is TOC accountability data.)

We are reasonably sure there is sufficient data available to determine sample vs laboratory vs equipment vs operator variance with respect to off line analyses of all types. Any data that has been obtained can be used as a beginning training set even though it is incomplete. Strong guidance to engineering design can be obtained.

B. CONCLUSION

THE ECLSS WATER RECLAMATION SYSTEM HAS THE POTENTIAL AS AN EXCELLENT DEMONSTRATION SYSTEM WHERE TQM CONCEPTS REDUCE COSTS AND IMPROVE QUALITY.

FIGURES

INLINE MULTIVARIATE ANALYSIS FLOW APPARATUS

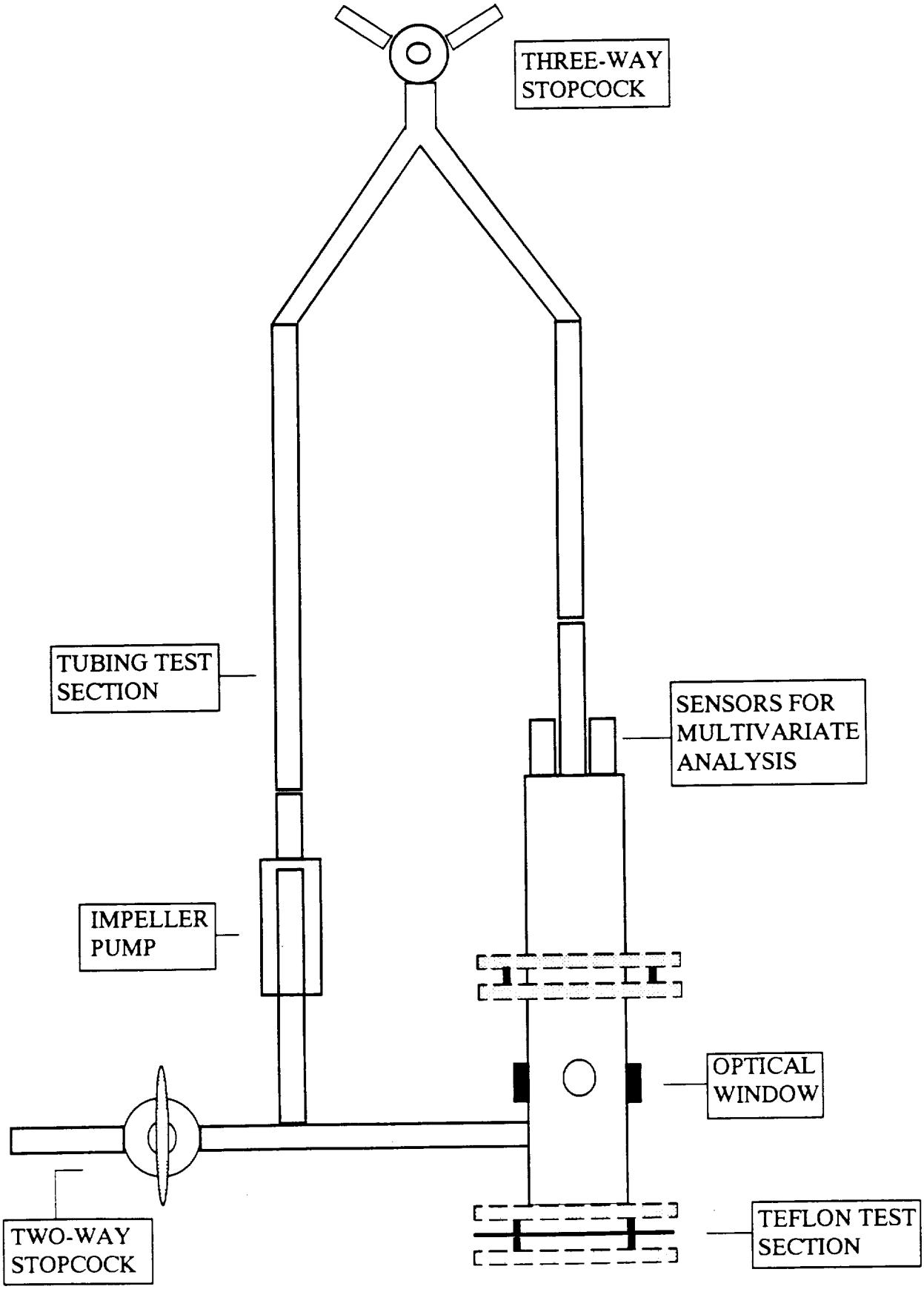


Figure 1

ATR NIR SENSOR FOR AIR OR AQUEOUS

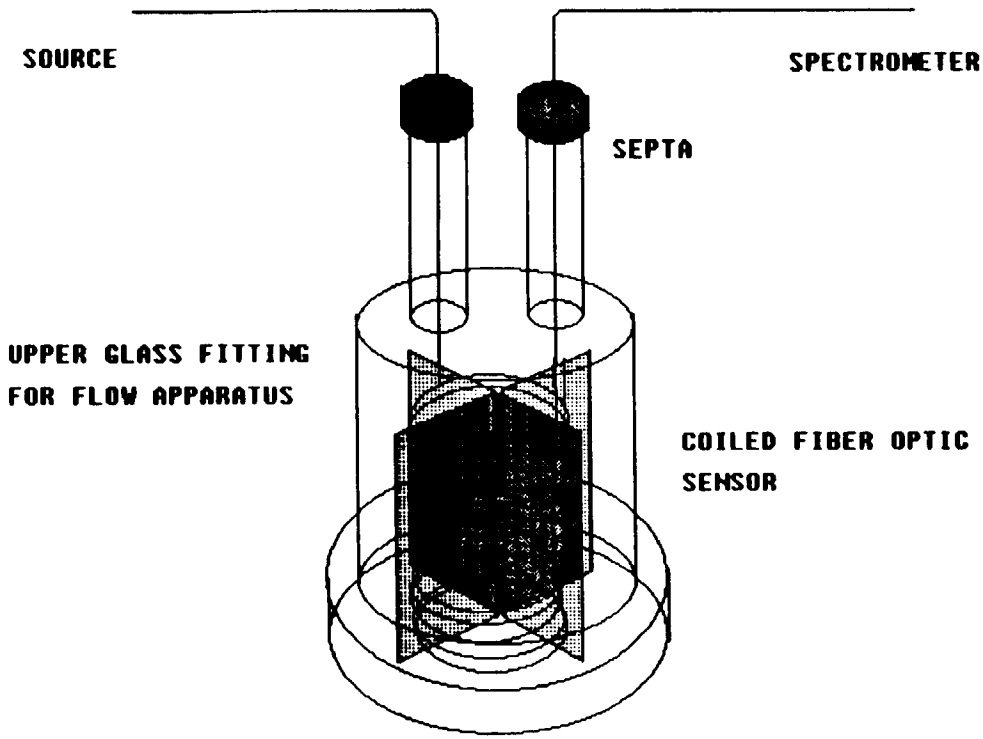


Figure 2

FIGURE 3

EFFECT OF PH ON I_3^- AND HIO

WITH PE = 10.51

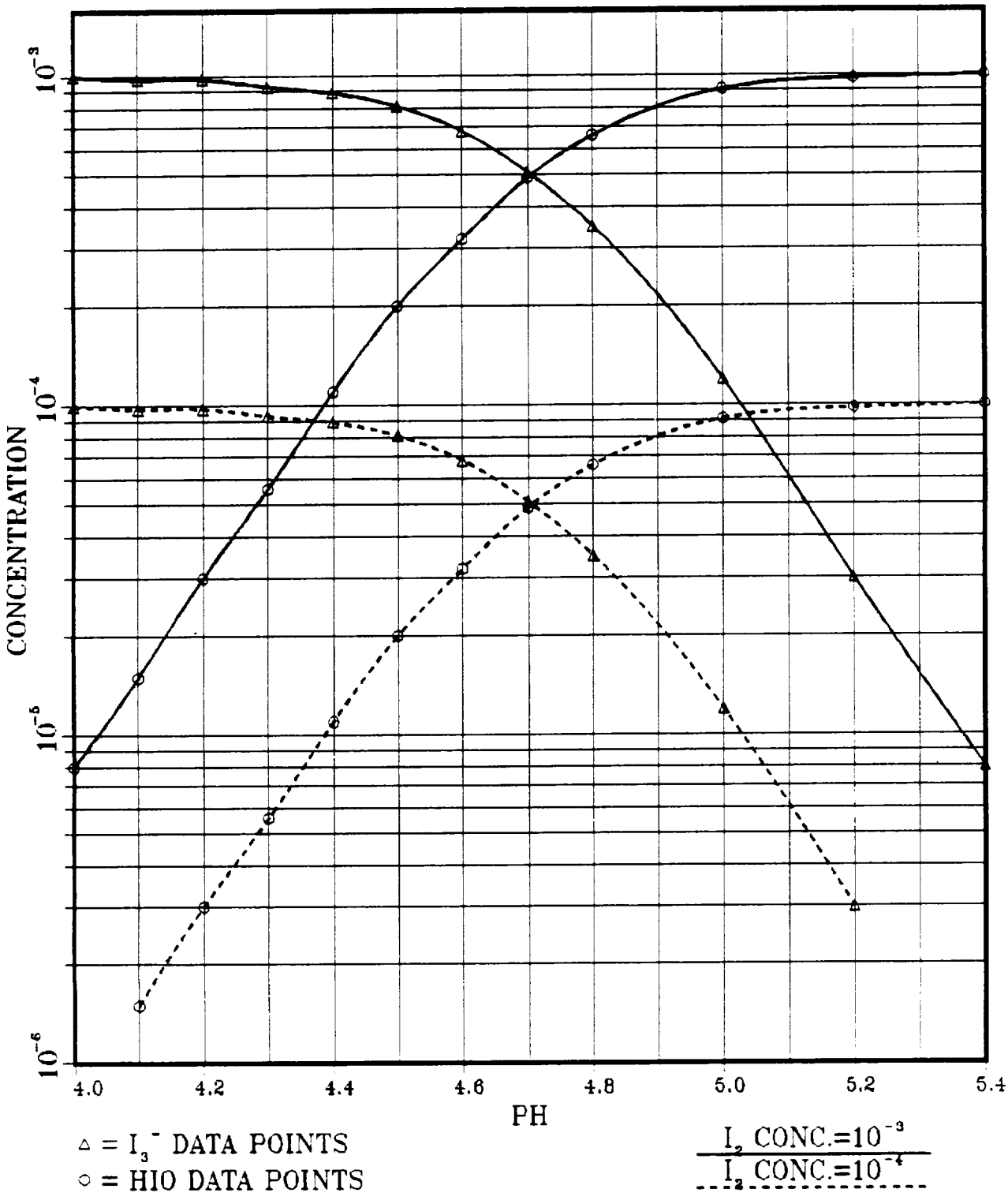
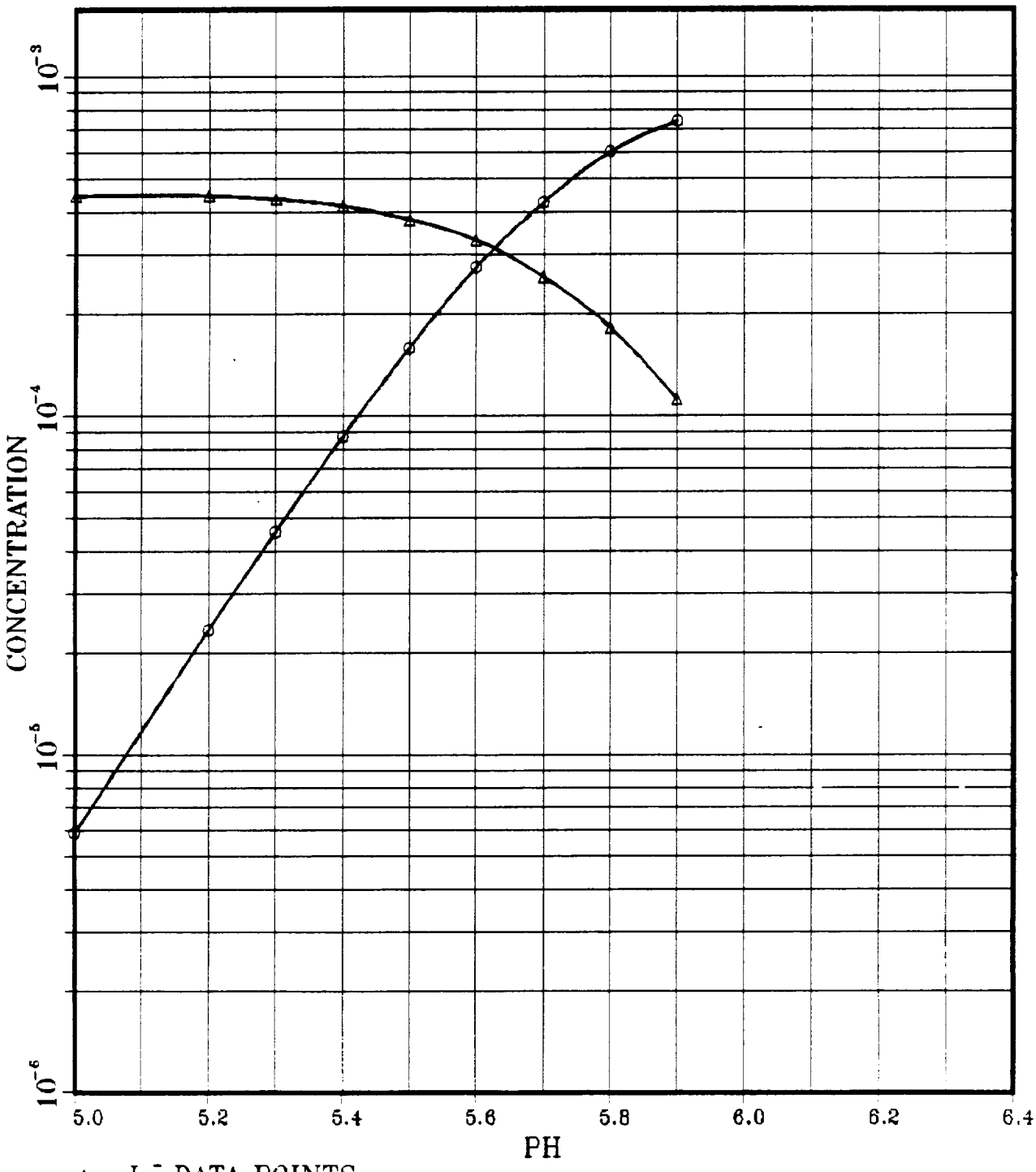


FIGURE 4

EFFECT OF PH ON I_3^- AND HIO

WITH PE=9.06



IODINE STANDADIZATION - I

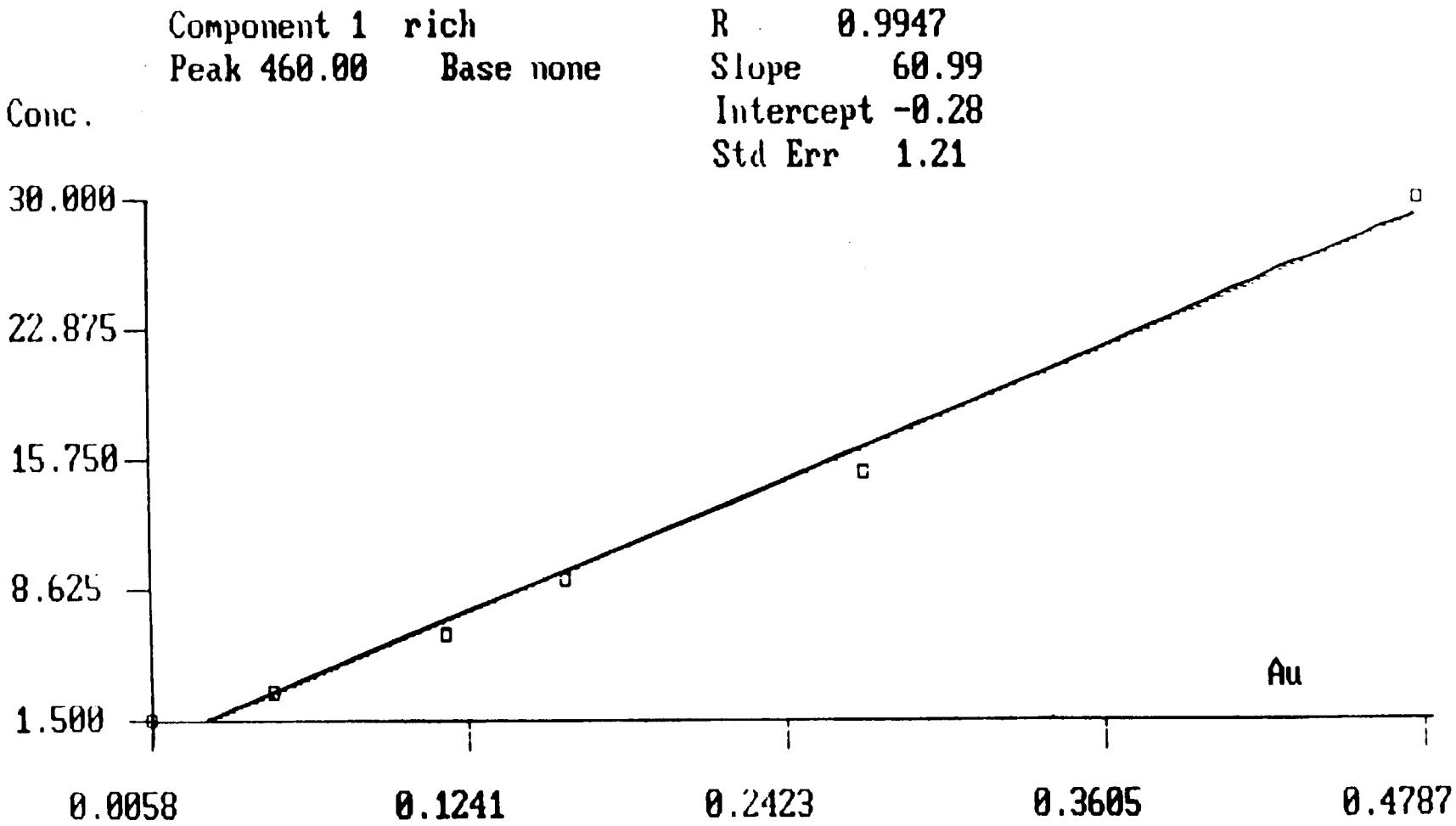


Figure 5

IODINE STANDARDIZATION -2

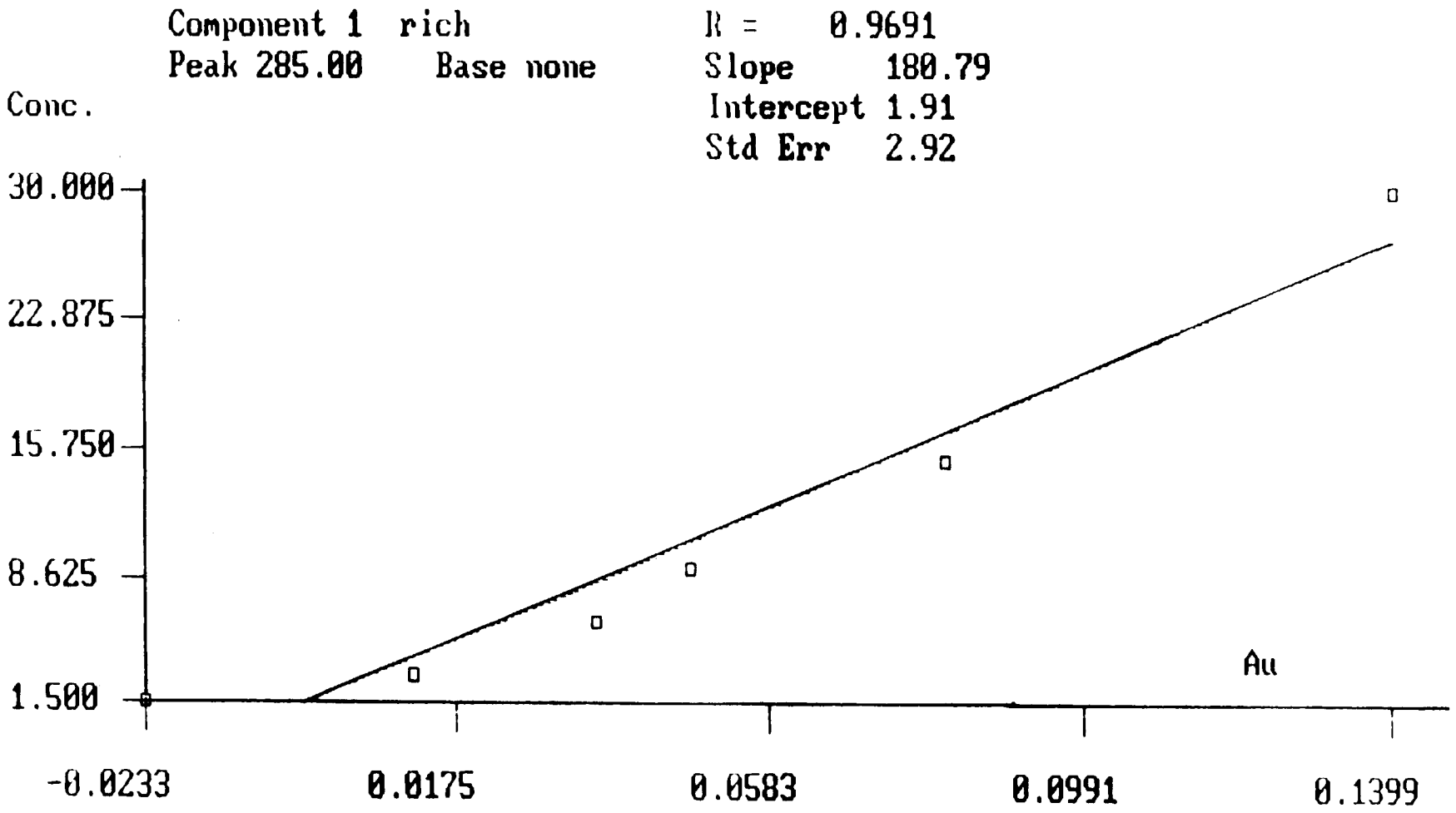


Figure 6

MODEL - ONE COMPOUND AT THREE WAVELENGTHS

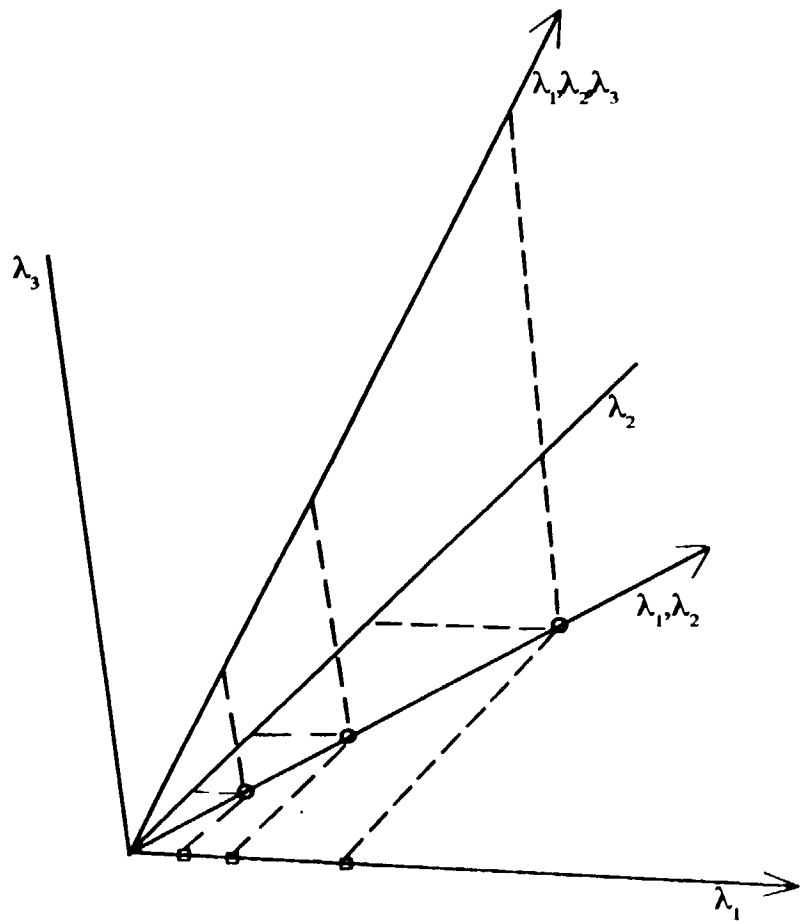


Figure 7.

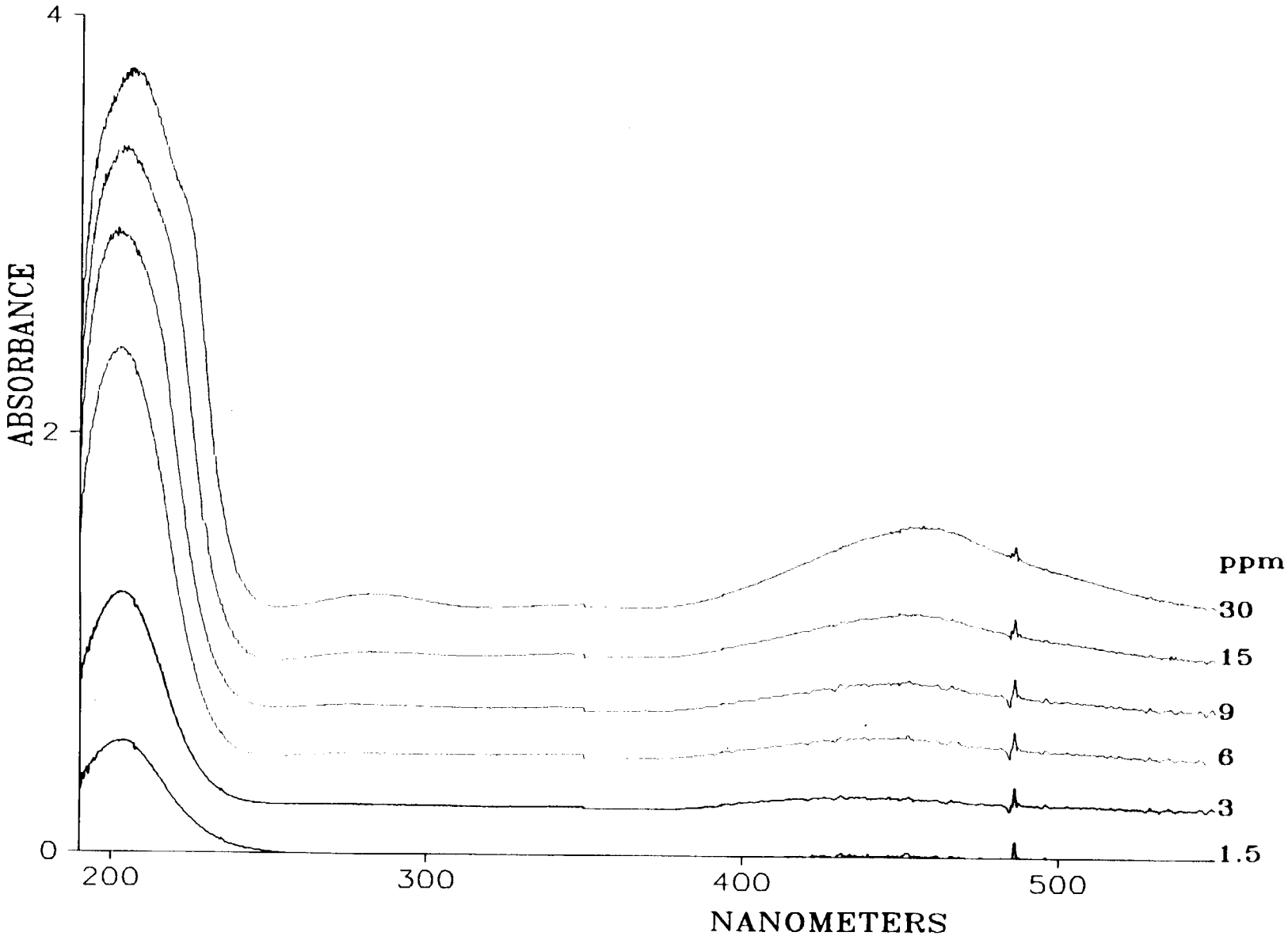


Figure 8. Iodine Standards

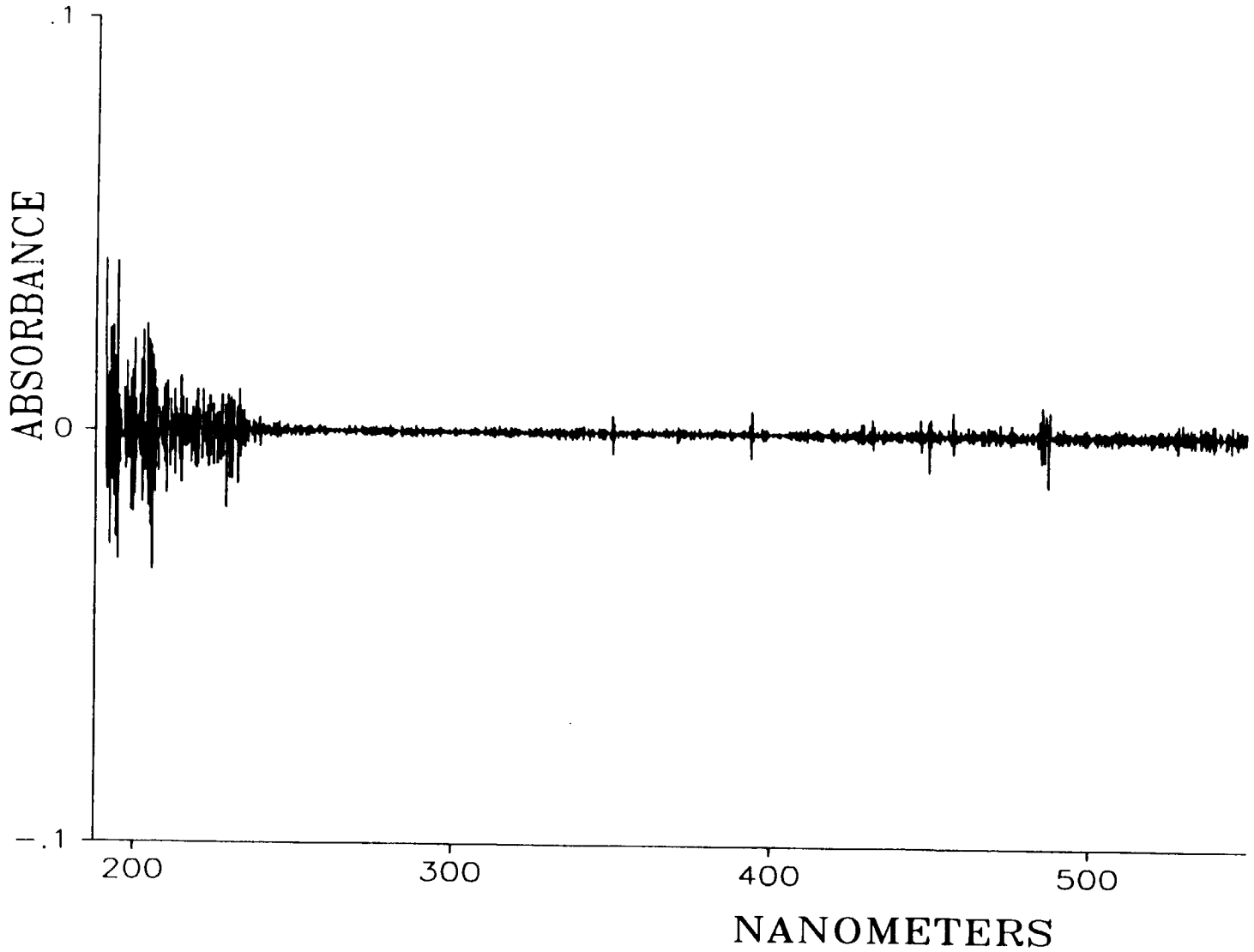


Figure 9. Difference Raw Data Minus Smoothed Data

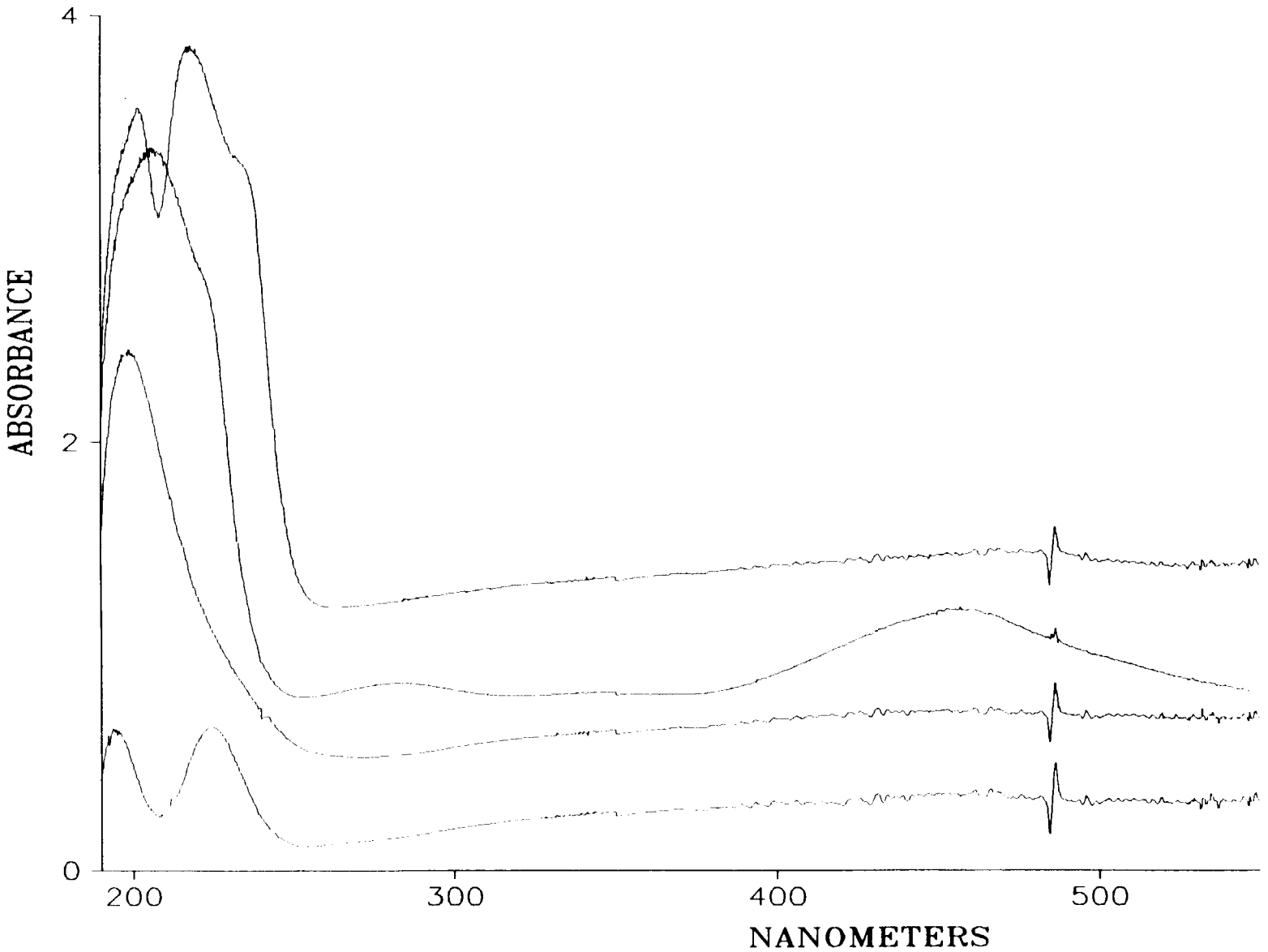


Figure 10. Spectra for Standards

TEFLON SECTION APPARATUS FOR IODINE VAPOR

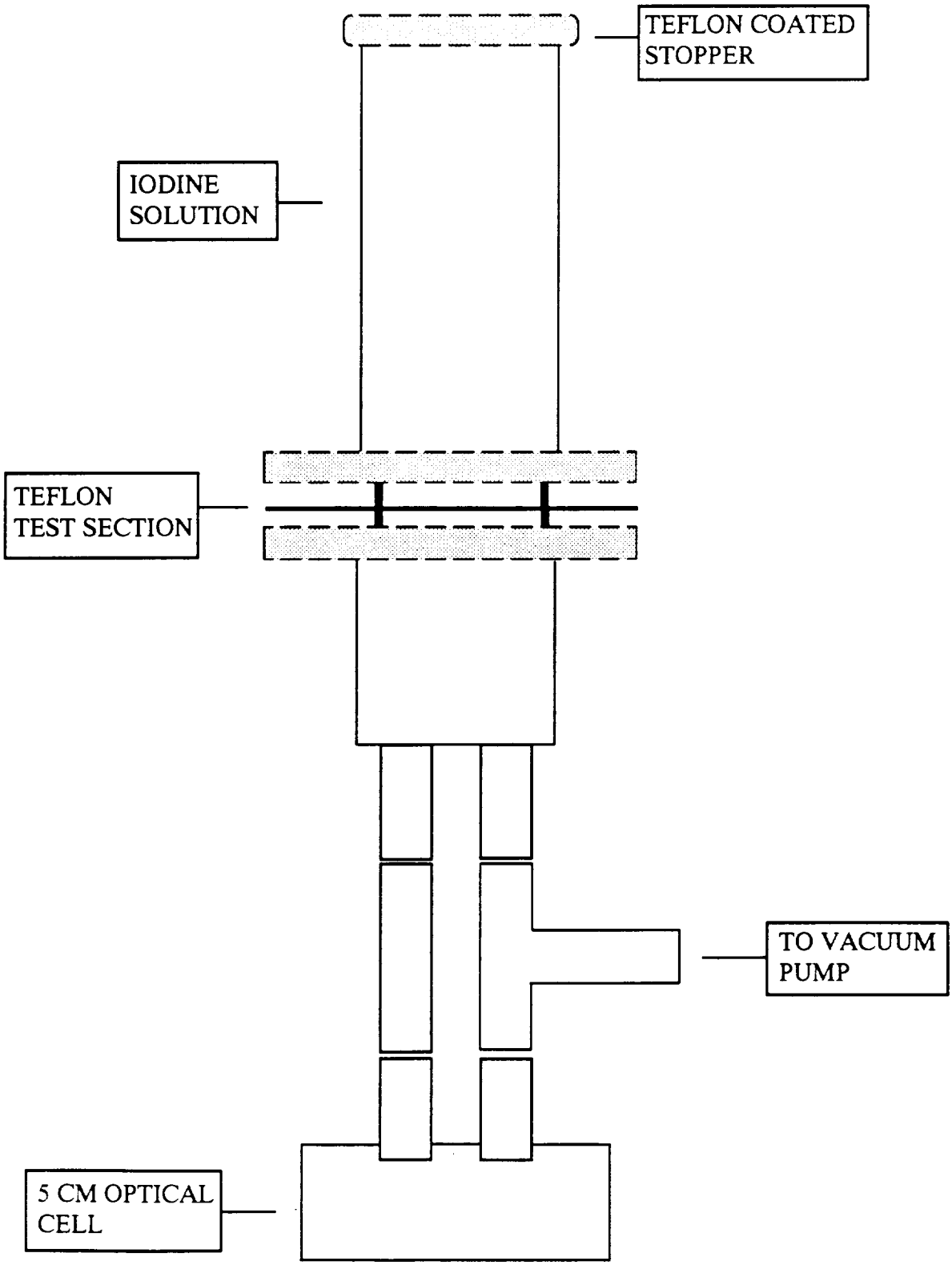


Figure 11.

CALIBRATION APPARATUS

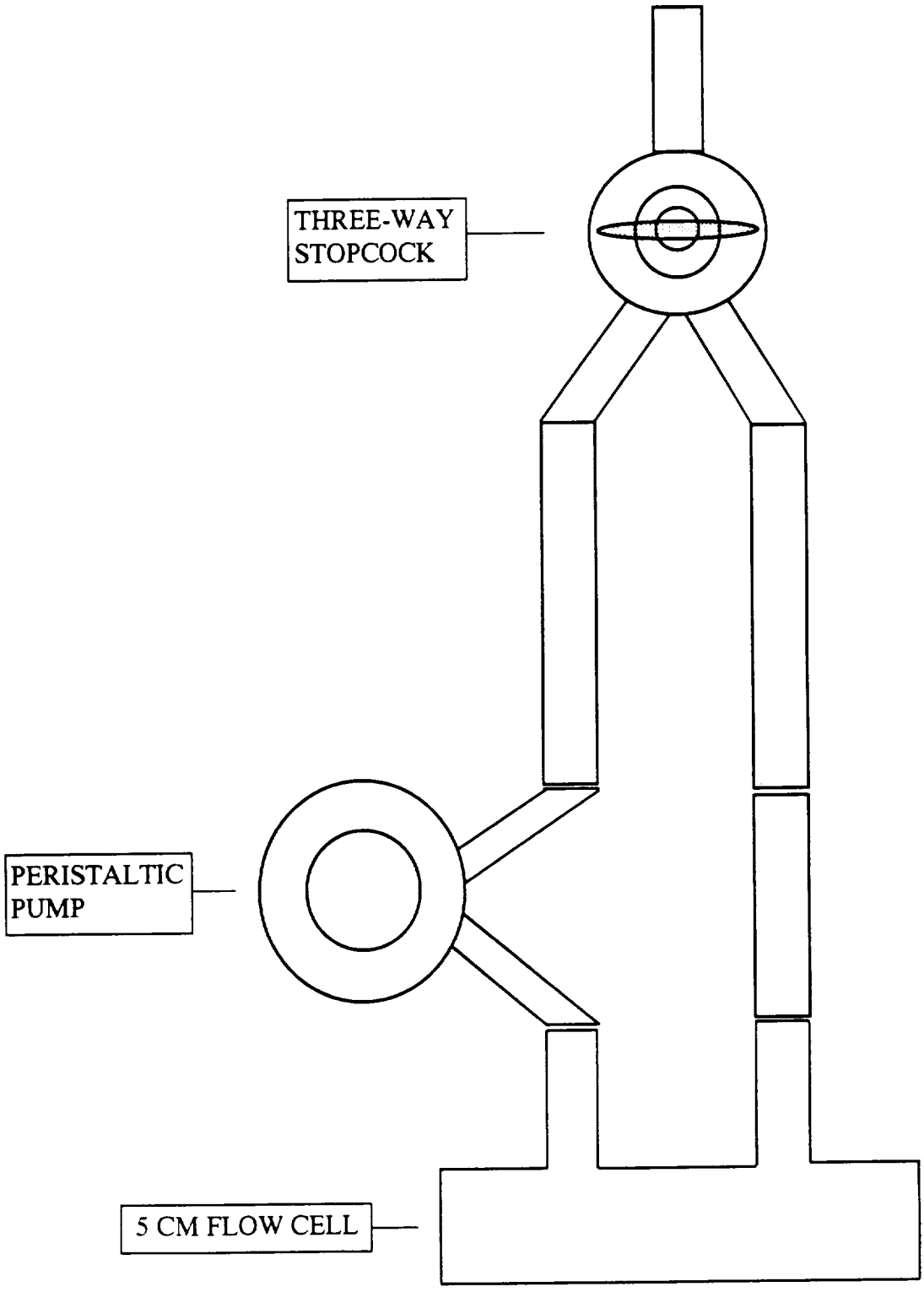


Figure 12.

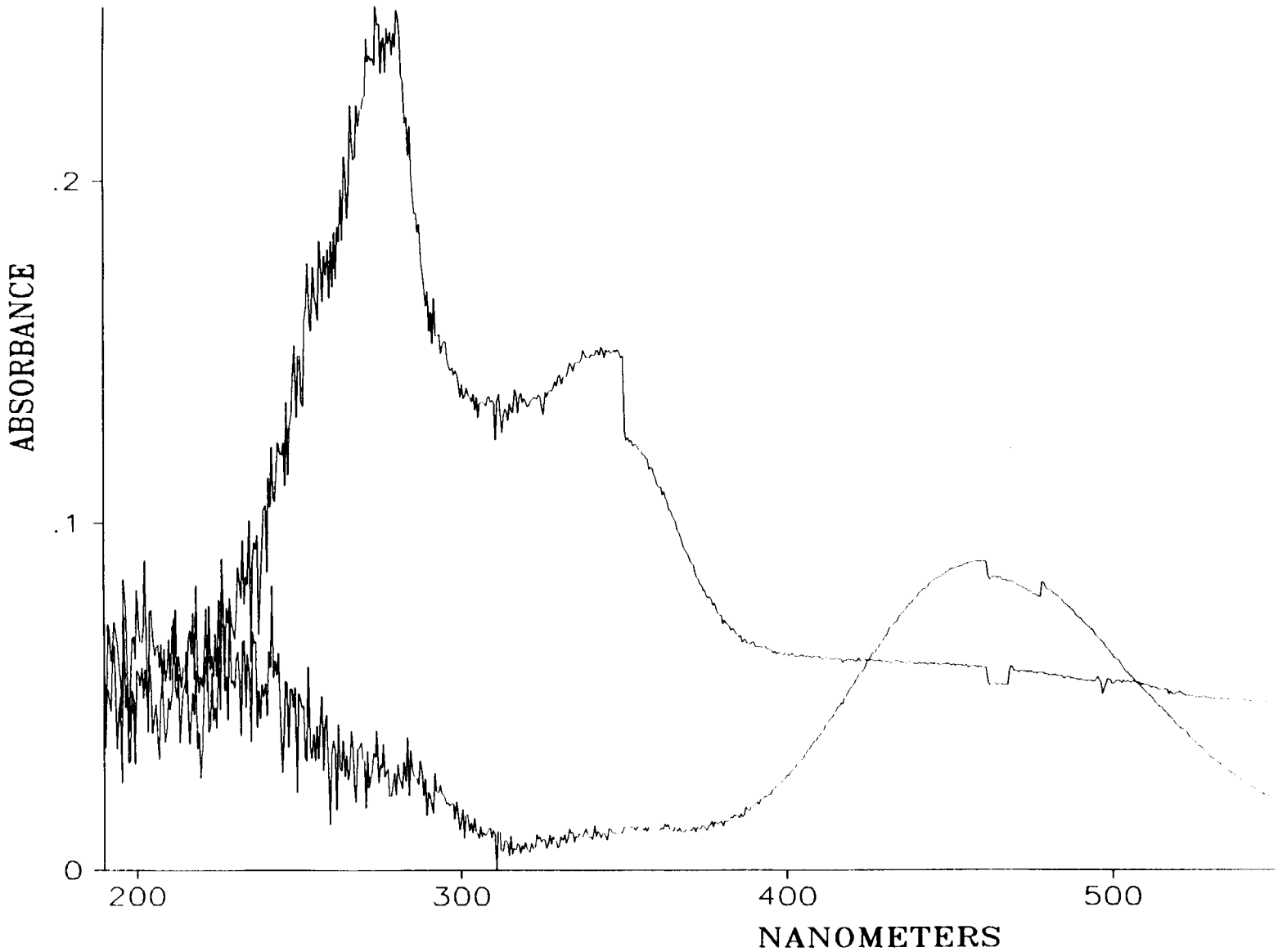


Figure 13. Spectrum From First Flow Experiment

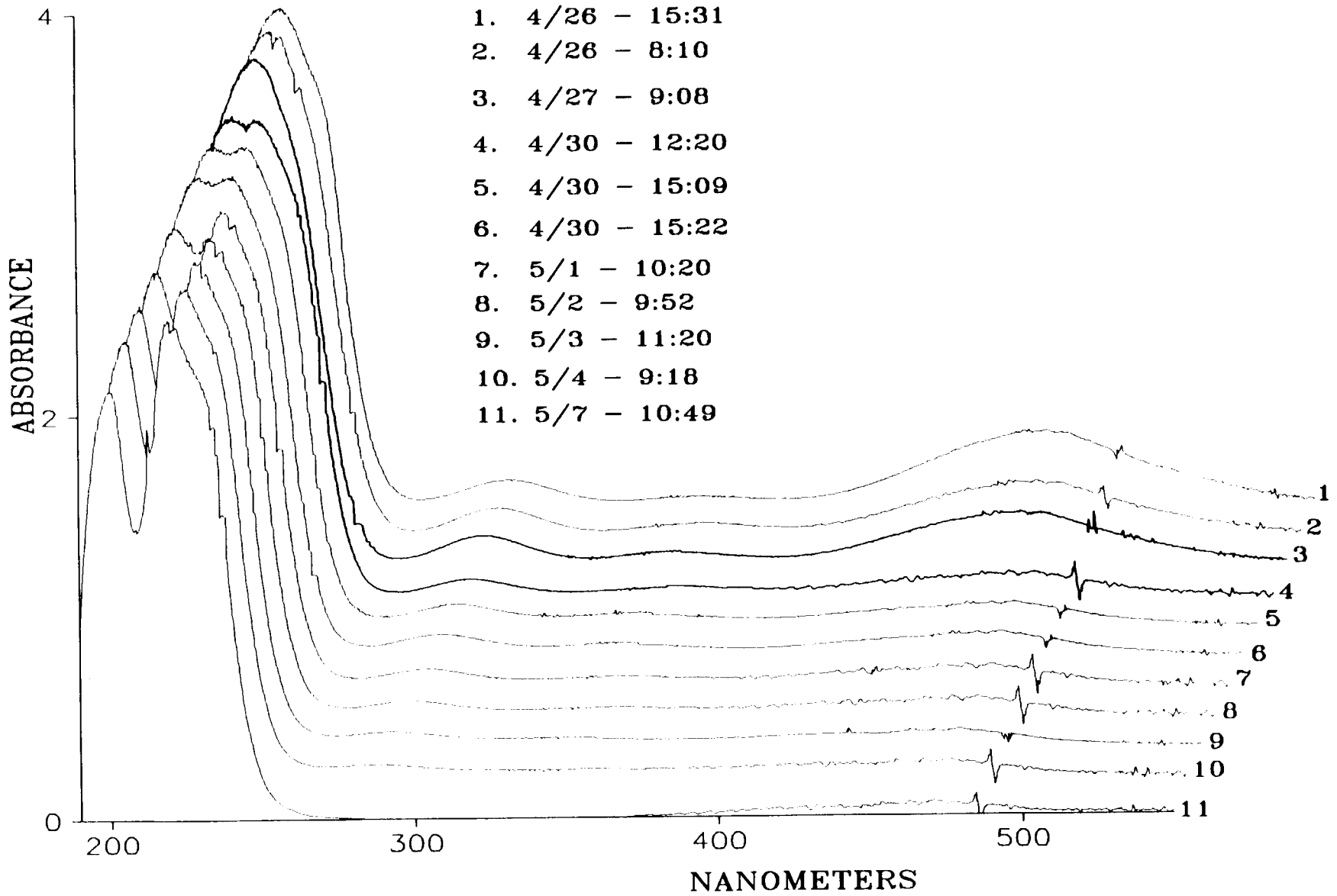


Figure 14. Iodine Stability

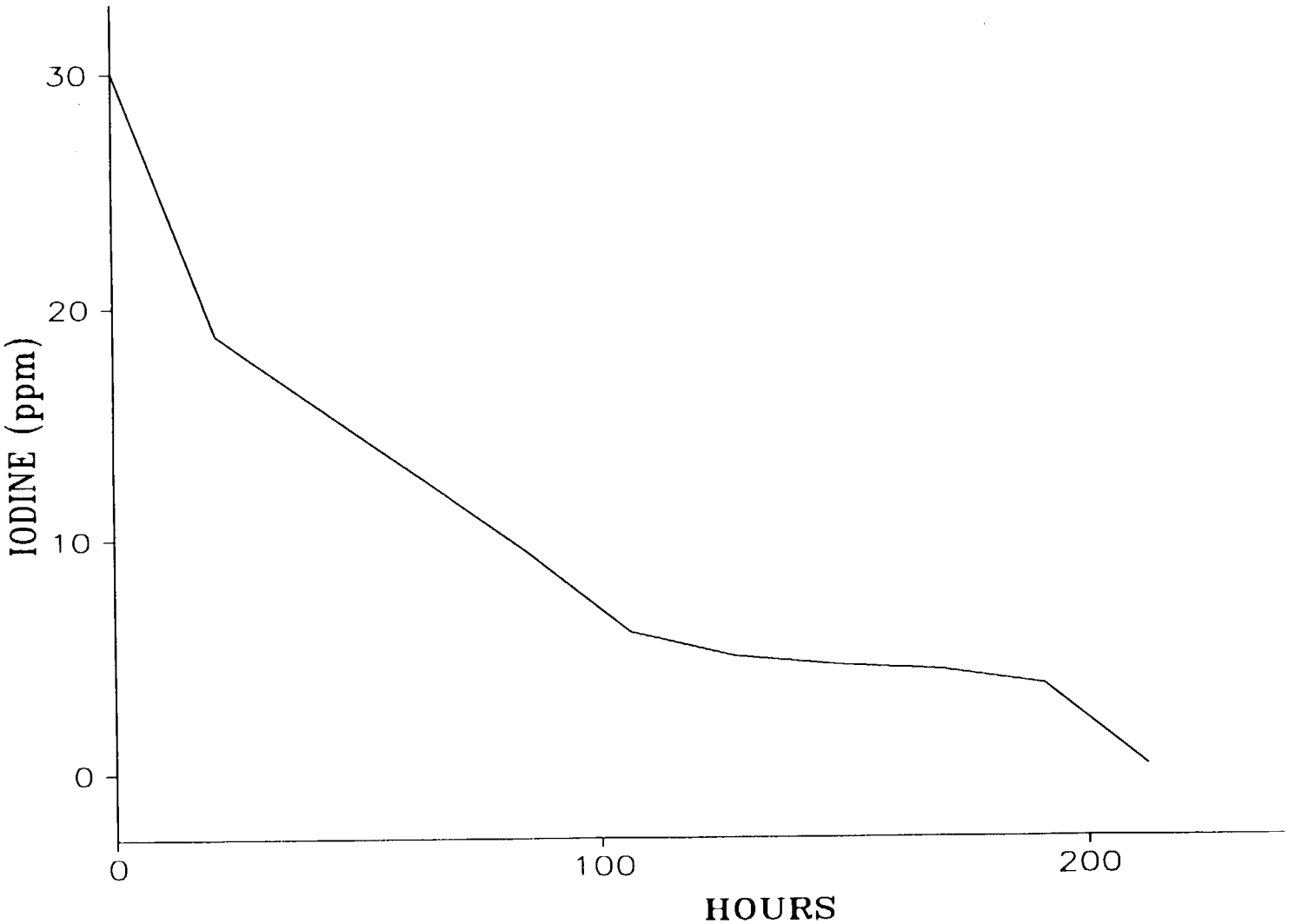


Figure 15. Iodine Loss in Circulating System

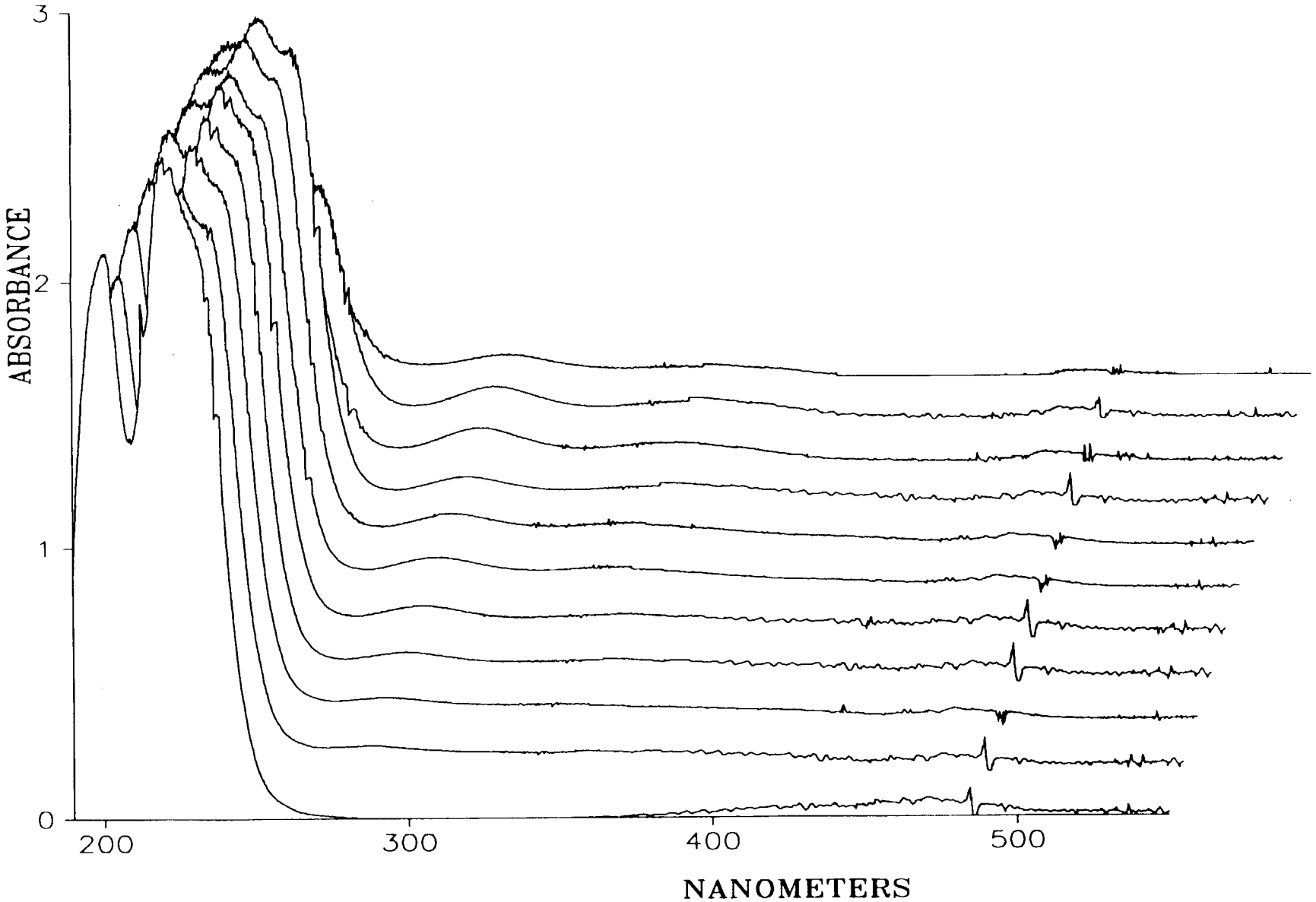


Figure 16. Residual After Removal of Iodine

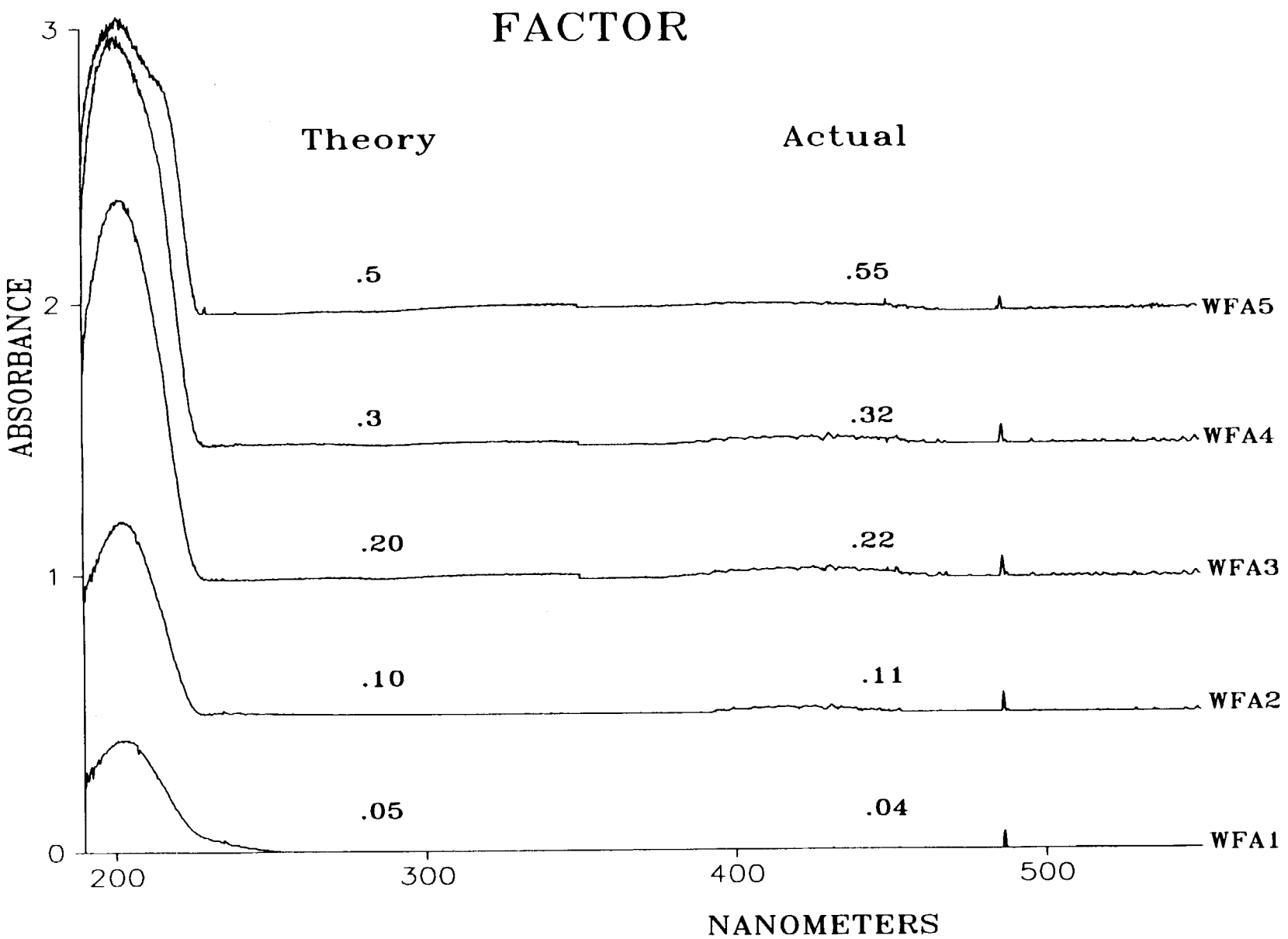


Figure 17. Difference Iodine – iodine 30 ppm

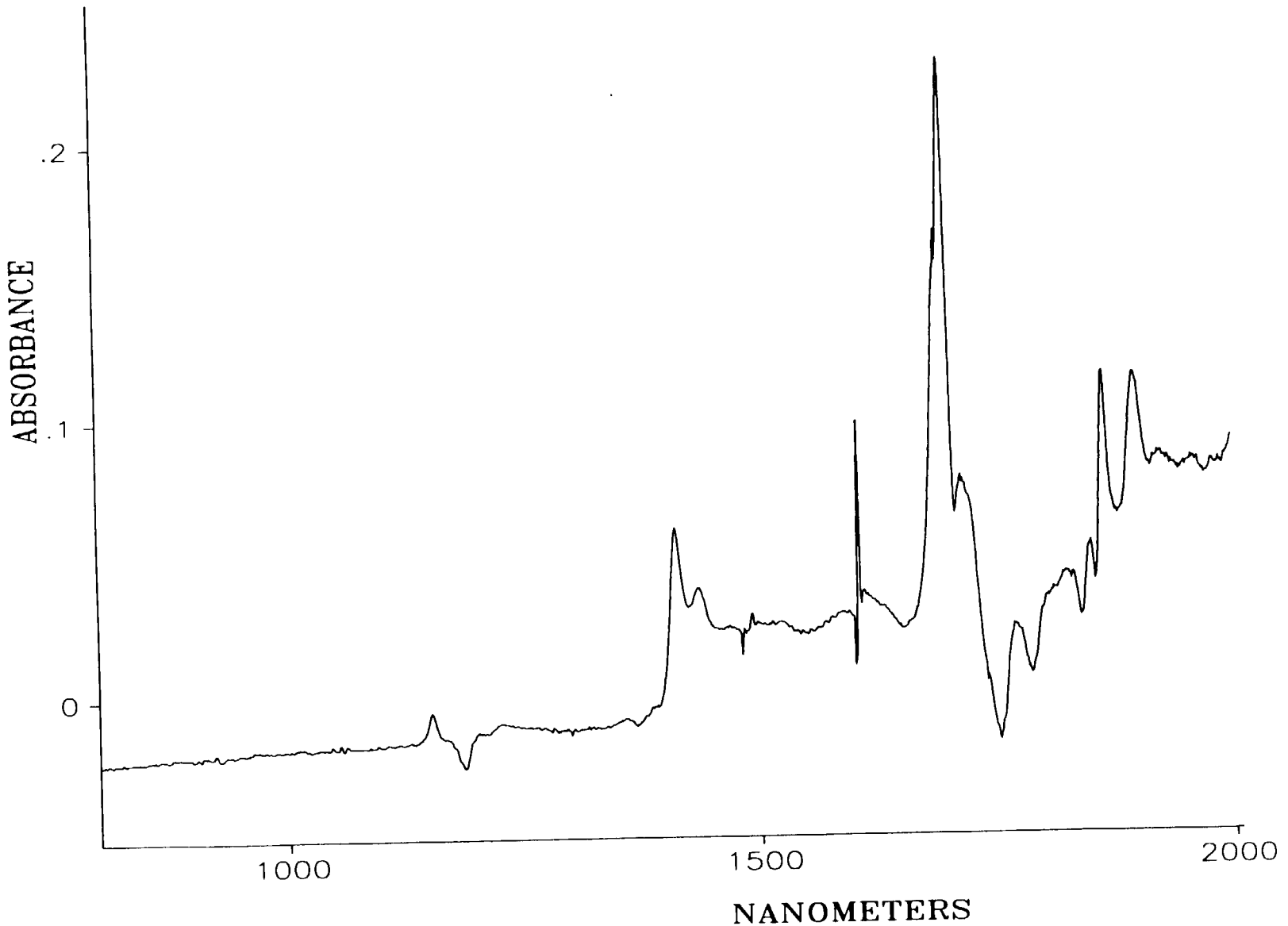


Figure 18. ATR Spectrum of 1:1 Ethanol and Chloroform

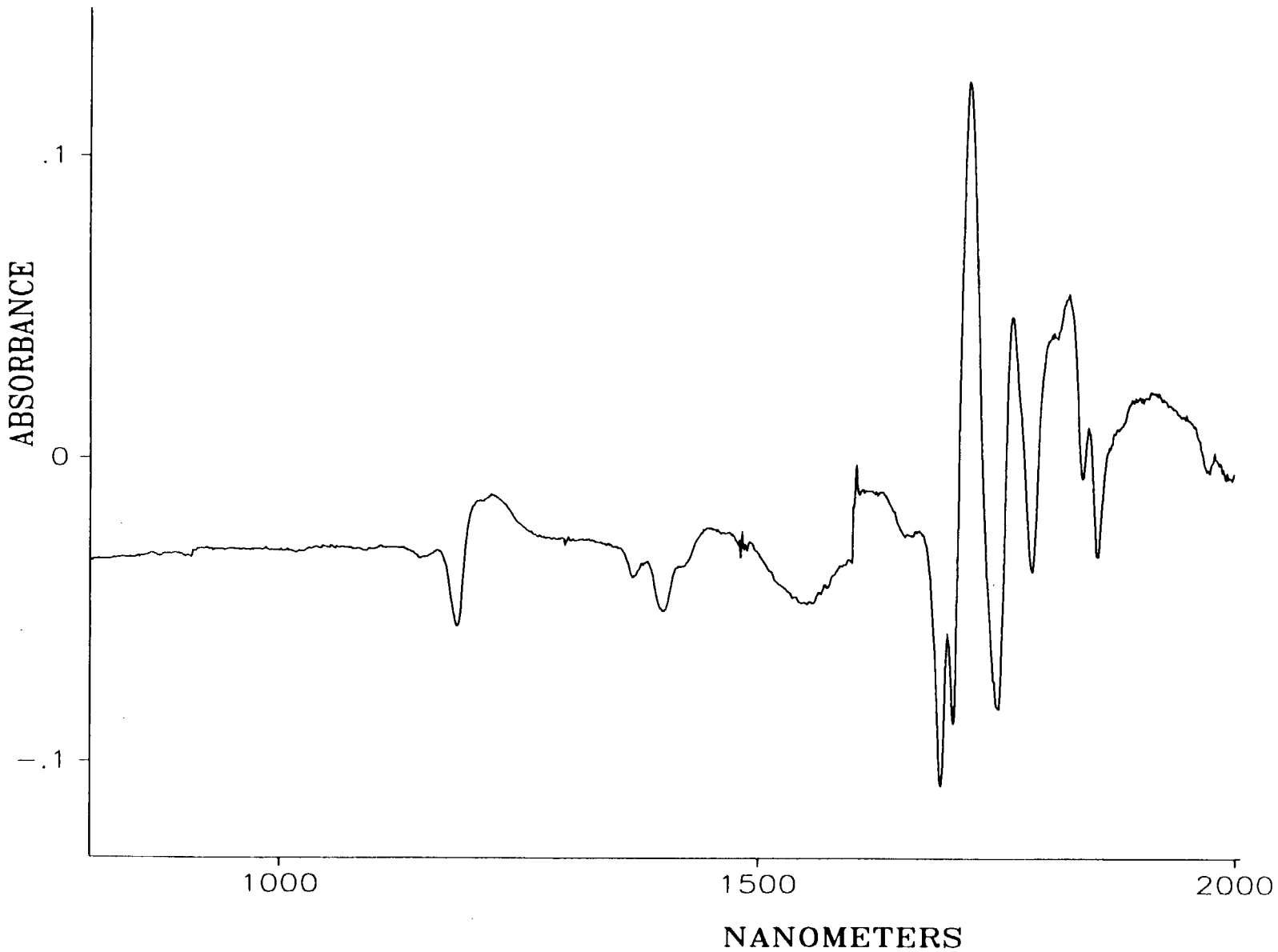


Figure 19. ATR Spectrum of Hexane

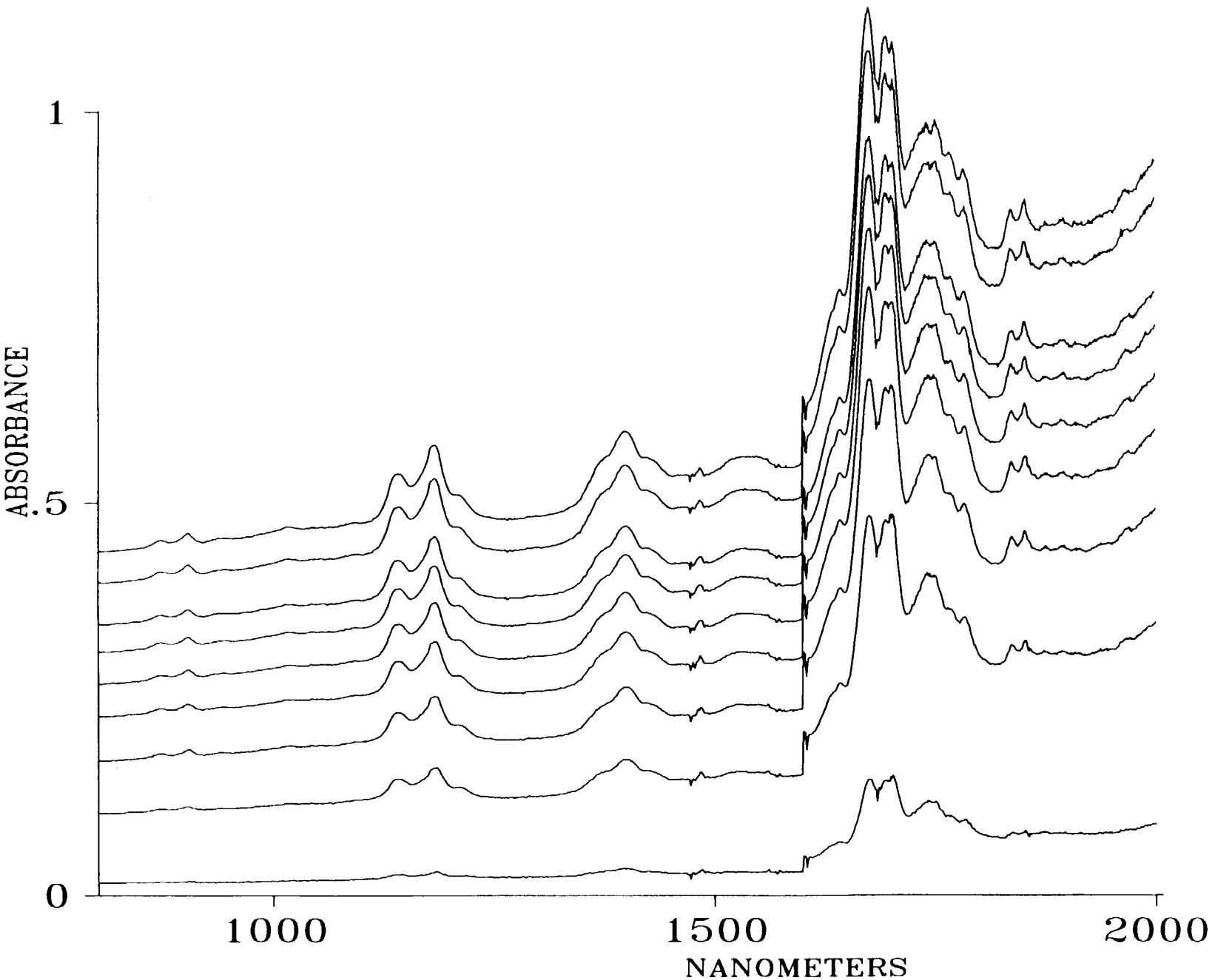


Figure 20. ATR Spectrum of Toluene Vapor in 30 Minute Intervals

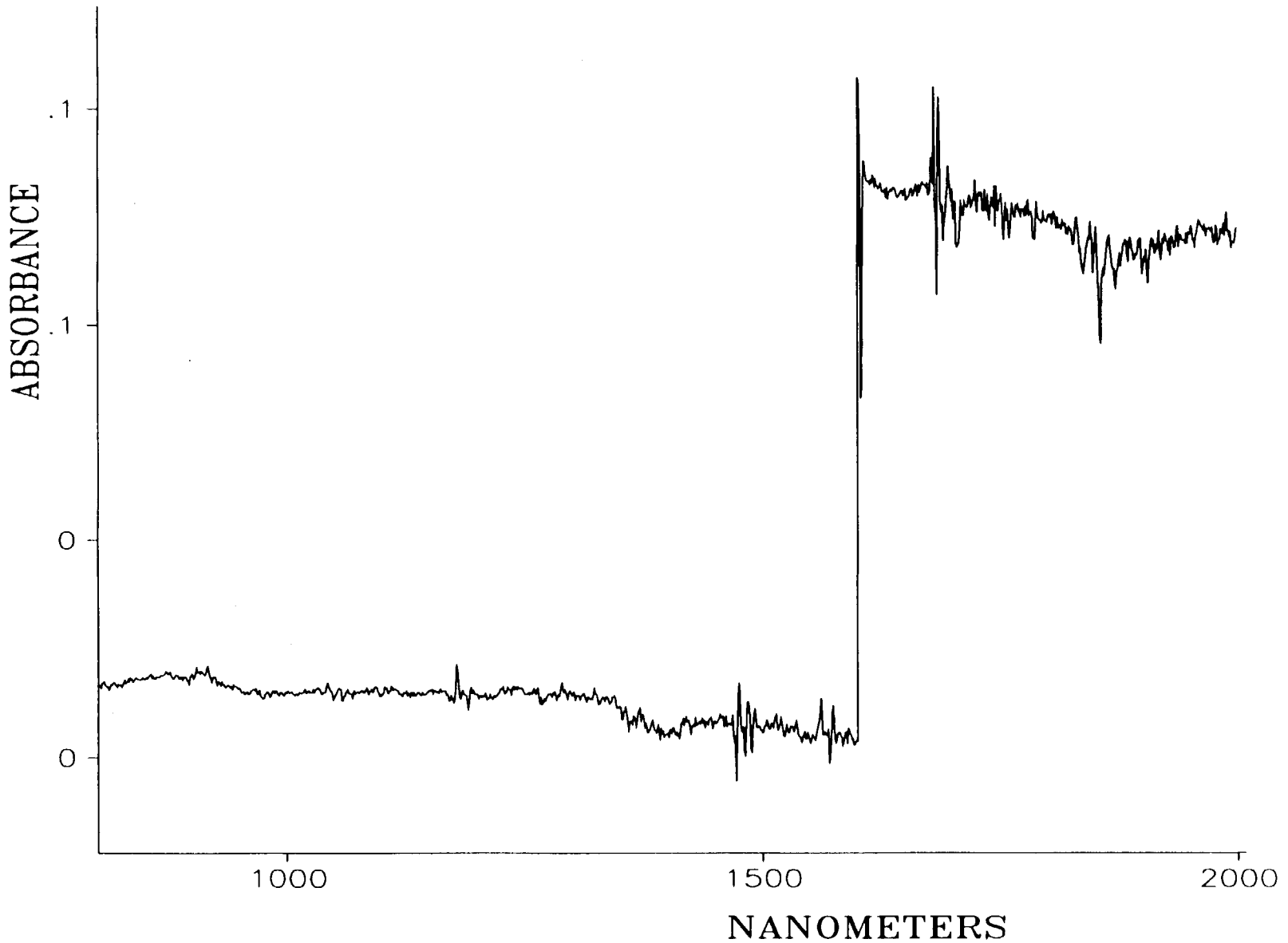


Figure 21. ATR Spectrum 15 minutes after removal of Toluene Solution

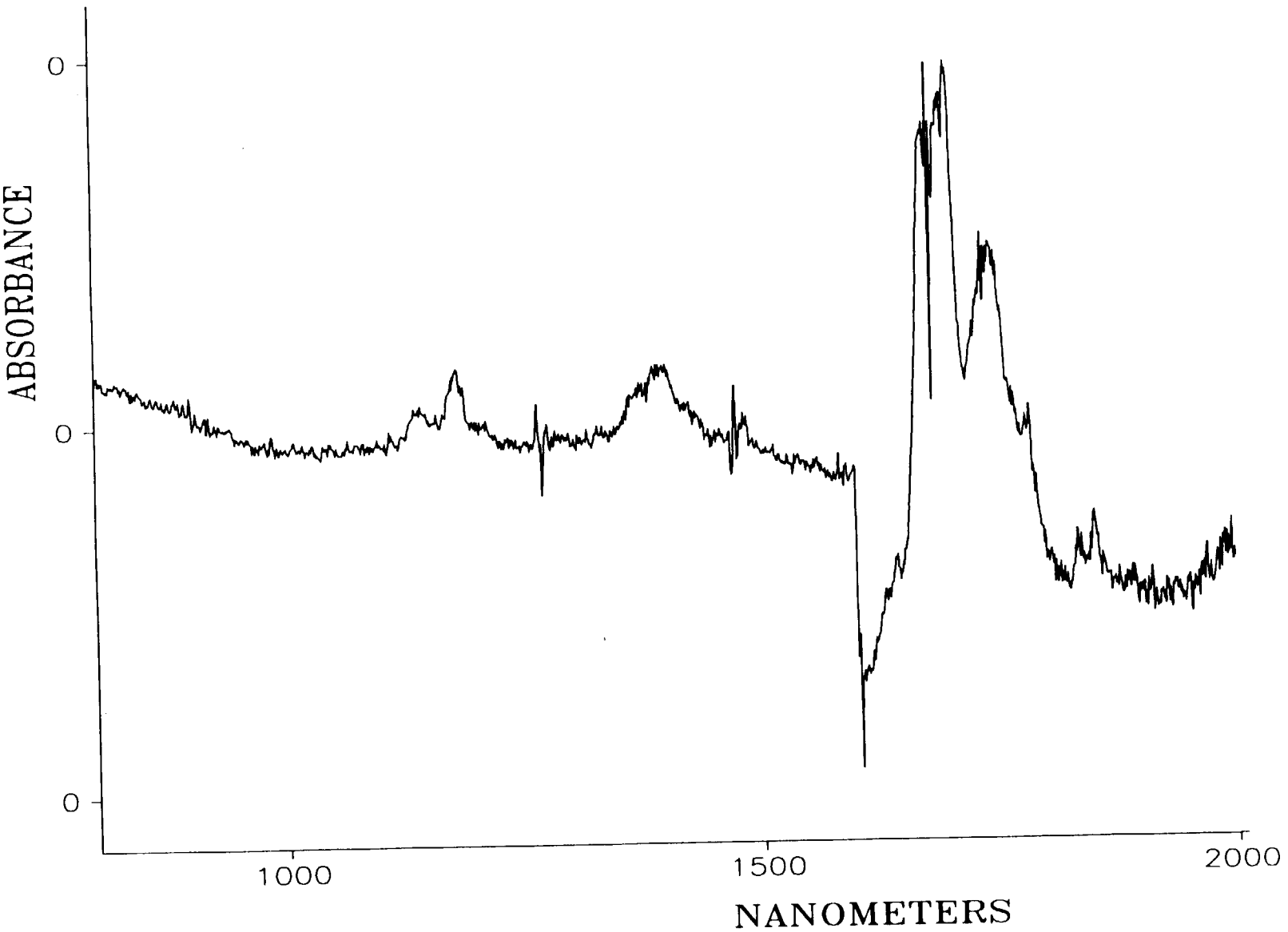


Figure 22. ATR Spectrum of Diluter Aqueous Toluene Solution

APPENDIX A

MODELS

Models may be divided into three general types: (1) quantitative or analytic, (2) qualitative, and (3) heuristic or empirical.

I. QUANTITATIVE OR ANALYTICAL.

A quantitative model consists of a mathematical representation of the dynamic features of a system as a set of differential equations, or a description in terms of state variables obtained from the knowledge of the fundamental physics and chemistry of the system. For systems with unpredictable behavior, stoichastic techniques can be employed. Shown at the top on Figure 1 are four quantitative models: (1) Most individuals are familiar with the equation representing an ideal gas. Data representing the state variables temperature, pressure, volume and mass can be correlated using the equation for an ideal gas. If any three of the properties are known the fourth can be calculated. For example, mass can be evaluated by measuring the volume, temperature, and pressure of a sample of gas. (2) Beer's Law is the principal equation used by spectroscopist. If the spectrum of each component of a mixture is known and the compounds in the mixture do not react, then the percent of each component is calculated by a manipulation of matricies. (3) Constructing a quantitative model for a system for measuring pH using a glass electrode or a hydrogen electrode is straight forward. The Nernst equation defines the relationship between the electropotential and the activity of the hydrogen ion. The Nernst equation also defines the change in potential with a change in temperature. Potential and current relate to the flow of electrons in conductors and are adequately described by Ohms law. Many successful pH meters have been developed and marketed using these equations to construct the model. As demonstrated by these examples, model building is a useful technique used by many scientists and engineers. This type of model is successful if the relationships are shown are linear and the variables can be separated. A first order model can be quickly constructed and instruments developed, manufactured and calibrated. A

MODELS (Y = f(X))

QUANTITATIVE OR ANALYTICAL

Ideal Gas Law $PV = nRT$

Beer's Law $A = ELC$

Nernst Equation $E = E^{o} - \frac{RT}{nF} \ln \frac{[Red]^{b}}{[Ox]^{a}}$

Tire Life (Diamond)

Life of Tire = f(type of Rubber) + f(Curing Temp) + f(Additives)
+ f(Groove Design) + f(Mixing Procedure) + f(Size of Tire)
+ f(Car Design) + f(Roadbeds) + f(Operating Speeds)
+ f(Atmospheric Conditions) + f(Operator Of Car)
+ f(Maintenance) + f(Type of Rubber)

$f(\text{Operating Speed}) = 87 \times EXP\left[7 - \left(\text{Speed} + \left(\frac{1}{\text{Speed}}\right)^{3}\right)\right]$

QUALITATIVE

A SYSTEM IS REPRESENTED OR DESCRIBED BY ITS COMPONENTS. CAUSAL RELATIONSHIPS ARE EMPHASIZED (FREQUENTLY USED IN EXPERT SYSTEMS AND NEURAL NETWORKS)

HEURISTIC OR EMPIRICAL

A SET OF EQUATIONS ABOUT A SYSTEM GENERATED THROUGH EXPERIMENTATION, STATISTICS, AND VALIDATION

Figure 1

closely related technique of model building is represented by the example taken from Diamond, William J., PRACTICAL EXPERIMENTAL DESIGNS. This example represents a synthesized linear model. It is assumed that the variables are separatable. Each of the individual functions is obtained by planning experiments where only one variable is changed at a time, that particular variables is modeled, and then the entire set of experiments is summed.

Quantitative of analytical models are sufficient for classical science and are particularly important to demonstrating causality. Science and technology has advanced our abilities to measure with greater precision, to see smaller objects, and to see objects at greater distance. Many examples are found where the variables interact and are not linear or separatable. Pure specimens of many of the chemicals in complex mixtures as found in nature are not available. Therefore, other methods of model building must be considered.

II. QUALITATIVE.

For a qualitative model a system is represented in terms of its components. The casual relationship between the components, corresponding to physical laws are emphasized. Functional descriptions of the components are utilized to generate rules and constraints for the entire system. This type of model can be demonstrated using the Ideal Gas Law. If the temperature and pressure of a volume of gas was being monitored in a container of fixed volume and a temperature rise is observed without an increase in pressure, as would be expected in a non-reacting system, a chemical reaction in which the number of molecules of gaseous product is less than the number of molecules of gaseous reactants may be postulated. This would not be a valid conclusion if a transducer malfunctions. The first rule therefore could be: if the above condition is observed check for the malfunction of a transducer. This type of model is frequently used for AI and expert system algorithms. Model of this type will be useful in describing monitoring system where detecting and quantitating a change in an analyte is more important than the base line composition. This type of model is less useful for modeling diagnostic systems.

III. HEURISTIC OR EMPIRICAL.

In a heuristic model a set of rules about the system is generated through experimentation and use of statistical algorithms. In the absence of a good theoretical model an empirical model may be the only effective method for the initial model of a system. Successful empirical models are derived through extensive testing. A major advance in the building of heuristic models resulted from the recognition that "the world" is multivariate and non linear. This has led to the development of new algorithms for use in multivariate analyses.

The amount of information obtained from an univariate statistical analysis is limited. For a single type of measurement on several samples, the samples can be represented on an axis or number line representing all possible values of the measurement. With univariate statistical techniques, some information can be learned about the one dimensional distribution of these points. More information can be gained when the values of a separate measurement are known for these samples. It is then possible to plot the points in a plane. As additional dimensions are added, more information is available.

The human being is the best pattern recognizer in two dimensions, or sometimes three dimensions. As additional types of measurements are made on the samples of a population, the potential to gain information about the properties of the sample increases, but the human ability to "see" patterns rapidly decreases. The computer, however, is able to manipulate points in multidimensional space, perform operation on them, and recognize the patterns in the higher order space. Cluster analysis and pattern recognition algorithms are designed to model systems and determine or predict properties of systems that are not themselves directly measurable but are related to measurements on the systems and give information that is sought. The pattern recognition approach is oriented to discover hidden relationships and to develop accurate and useful prediction and classification

models for aspects (properties of categories) that are difficult or impossible to measure, using features or variable that are easy and inexpensive to measure and a training set of knowns. As each individual has a characteristic finger print, each instrument, each transducer, and each sample has a characteristic data vector. If the data vectors for two instruments are the same, the instruments are identical. If the data is different for identical samples by several operators then the instruments or operators are responsible. If the data vectors are not identical, they can often be separated into two or more vectors that can be identified. This type of modeling is particularly useful for evaluating data taken using several instruments and several operators to identify malfunctioning instruments and/or unqualified technicians.

Empirical models have been used beneficially for near infrared analyses (NIRA). Data vectors (spectra) are taken for samples of known composition. An empirical model is established based on the variance covariance matrix. The content of unknown samples can then be quantitated (predicted). These models have been used extensively for analyses of agricultural products,drugs, and textiles.

Relationship between error of prediction and complexity of the calibration are shown on Figure 2. As indicated if only one observable is used, the error of prediction caused by interferants is high. As longer data vectors are used the error of prediction decreases. An example from spectroscopy is the use of a single wavelength and Beer's law vs the entire spectrum and Beer's law. When a factor analysis approach is used, the use of too many factors results in including noise in the data. The training set must match the experimental conditions.

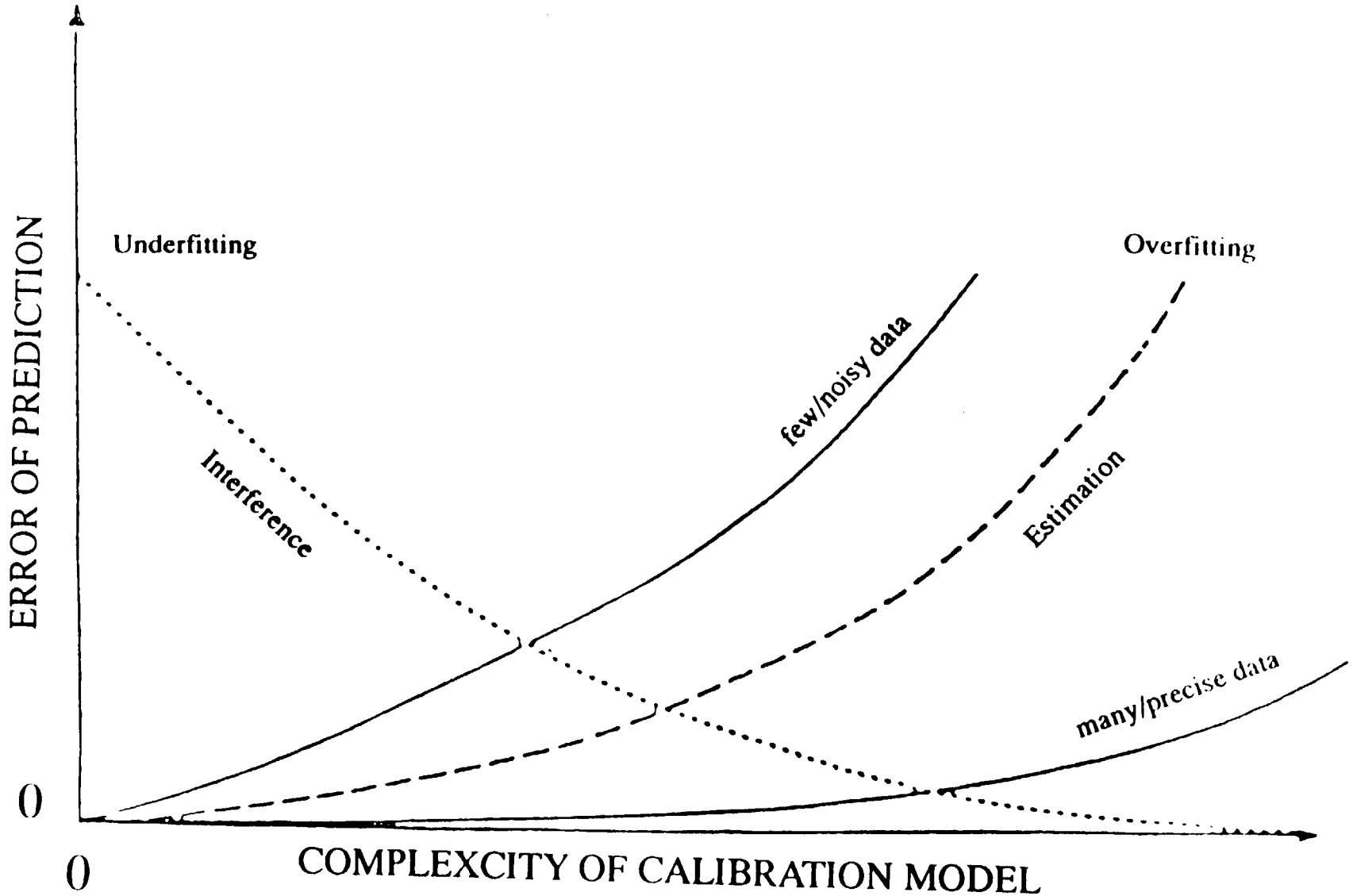


Figure 2. THE EFFECTS OF INCREASED AMOUNT OF DATA AND BETTER DATA